jD - /b -

$Box SEQ/A$

Attorney Docket No.	24736-2033
First named inventor	Andreas Braun
Express mail label #	EL406143913US
Date of mailing	October 13, 2000

TRANSMITTAL OF UTILITY APPLICATION UNDER 37 C.F.R. §1.53

Application Elements	Accompanying Application Papers

Application Elements

1. [X] Fee Transmittal Form

2. [X] Specification containing <u>135</u> pages (including claims and Abstract), a Sequence Listing (110 pages).

   a. Title: **METHODS FOR GENERATING DATABASES AND DATABASES FOR IDENTIFYING POLYMORPHIC GENETIC MARKERS**

   b. Number of claims: <u>100</u>

3. [X] 49 sheets of drawings with 58 Figures

4. [ ] Declaration

5. [X] Sequence Listing

   [X] Paper copy (identical to computer copy)

   [X] Computer readable copy

   [ ] Verified statement

Accompanying Application Papers

6. [ ] Assignments

7. [X] Small Entity Status is claimed

8. [ ] Preliminary Amendment

9. [X] Return Receipt Postcard

SIGNATURE OF ATTORNEY/AGENT

HELLER EHRMAN WHITE & McAULIFFE LLP

Stephanie Seidman
Registration Number: 33,779

[X] Benefit of priority under 35 U.S.C. §119(e) claimed to U.S. provisional application Serial Nos. 60/217,658, 60/159,176 and 60/217,251. Benefit of priority under 35 U.S.C. §120 is also claimed. This application is a continuation-in-part of U.S. application Serial No. 09/663,938, filed September 19, 2000. The application and provisional applications are incorporated by reference in their entirety

CORRESPONDENCE ADDRESS

NAME	Stephanie Seidman Registration No. 33,779 Heller Ehrman White & McAuliffe LLP
Address	4250 Executive Square, 7th Floor, La Jolla, CA 92037
	Telephone: 858.450-8400	Facsimile: 858.587-5360

FEE TRANSMITTAL ACCOMPANYING UTILITY APPLICATION UNDER 37 C.F.R. §1.53	Attorney Docket No.	24736-2033
	First named inventor	Andreas Braun
	Express mail label #	EL406143913US
	Date of mailing	October 13, 2000

FEE CALCULATION FOR CLAIMS AS AMENDED

a) Basic Fee $ 710.00
b) Independent Claims 15 - 3 = 12 x $ 80.00 $ 960.00
c) Total Claims   100 - 20 = 80 x $ 18.00 $1440.00
d) Fee for Multiple Dependent Claims - $270.00 $ 0.00
                              TOTAL FILING FEE $3110.00

[X] Status as Small Entity is claimed
     reducing Fee by one-half to $1555.00

[X] A check in the amount of $1550.00 to cover the fee for filing the application.

[ ] Charge $ .00 to Deposit Account No. 50-1213

[X] The Commissioner is hereby authorized to charge any fees, including the filing fee and excess claims fee, that may be required in this application under 37 C.F.R. §§ 1.16-1.17 during its entire pendency, or credit any overpayment, to Deposit Account No. 50-1213. If proper payment is not enclosed, such as a check in the wrong amount, unsigned, post-dated, otherwise improper or informal, or absent, the Commissioner is authorized to charge the unpaid amount to Deposit Account No. 50-1213 during the entire pendency of this application. This sheet is filed in duplicate.

CORRESPONDENCE ADDRESS				
NAME	Stephanie Seidman Registration No. 33,779 Heller Ehrman White & McAuliffe LLP			
Address	4250 Executive Square, 7th Floor, La Jolla, CA 92037			
	Telephone: 858.450.8400		Facsimile: 858.587.5360	
Submitted by:				
Typed or printed name	Stephanie Seidman		Reg. Number	33,779
Signature		Date 10/13/00	Deposit Account	50-1213

METHODS FOR GENERATING DATABASES AND DATABASES FOR IDENTIFYING POLYMORPHIC GENETIC MARKERS

RELATED APPLICATIONS

Benefit of priority under 35 U.S.C. §119(e) to the following

5   provisional applications is claimed herein:

U.S. provisional application Serial No. 60/217,658 to Andreas Braun, Hubert Koster; Dirk Van den Boom, filed July 10, entitled "METHODS FOR GENERATING DATABASES AND DATABASES FOR IDENTIFYING POLYMORPHIC GENETIC MARKERS"; U.S. provisional application Serial

10   No. 60/159,176 to Andreas Braun, Hubert Koster, Dirk Van den Boom, filed October 13, 1999, entitled "METHODS FOR GENERATING DATABASES AND DATABASES FOR IDENTIFYING POLYMORPHIC GENETIC MARKERS"; U.S. provisional application Serial No. 60/217,251, filed July 10, 2000, to Andreas Braun, entitled "POLYMORPHIC KINASE

15   ANCHOR PROTEIN GENE SEQUENCES, POLYMORPHIC KINASE ANCHOR PROTEINS AND METHODS OF DETECTING POLYMORPHIC KINASE ANCHOR PROTEINS AND NUCLEIC ACIDS ENCODING THE SAME". This application is also a continuation-in-part of U.S. application Serial No. 09/663,968, to Ping Yip, filed September 19, 2000, entitled "METHOD

20   AND DEVICE FOR IDENTIFYING A BIOLOGICAL SAMPLE."

The above-noted applications and provisional applications are incorporated by reference in their entirety.

FIELD OF THE INVENTION

Process and methods for creating a database of genomic samples

25   from healthy human donors. Methods that use the database to identify and correlate with polymorphic genetic markers and other markers with diseases and conditions are provided.

BACKGROUND

Diseases in all organisms have a genetic component, whether

30   inherited or resulting from the body's response to environmental stresses, such as viruses and toxins. The ultimate goal of ongoing genomic

-1-

research is to use this information to develop new ways to identify, treat and potentially cure these diseases. The first step has been to screen disease tissue and identify genomic changes at the level of individual samples. The identification of these "disease" markers has then fueled

5 the development and commercialization of diagnostic tests that detect these errant genes or polymorphisms. With the increasing numbers of genetic markers, including single nucleotide polymorphisms (SNPs), microsatellites, tandem repeats, newly mapped introns and exons, the challenge to the medical and pharmaceutical communities is to identify

10 genotypes which not only identify the disease but also follow the progression of the disease and are predictive of an organism's response to treatment.

Currently the pharmaceutical and biotechnology industries find a disease and then attempt to determine the genomic basis for the disease.

15 This approach is time consuming and expensive and in many cases involves the investigator guessing as to what pathways might be involved in the disease.

Genomics

Presently the two main strategies employed in analyzing the

20 available genomic information are the technology driven reverse genetics brute force strategy and the knowledge-based pathway oriented forward genetics strategy. The brute force approach yields large databases of sequence information but little information about the medical or other uses of the sequence information. Hence this strategy yields intangible

25 products of questionable value. The knowledge-based strategy yields small databases that contain a lot of information about medical uses of particular DNA sequences and other products in the pathway and yield tangible products with a high value.

Polymorphisms

Polymorphisms have been known since 1901 with the identification of blood types. In the 1950's they were identified on the level of proteins using large population genetic studies. In the 1980's and 1990's many

5 of the known protein polymorphisms were correlated with genetic loci on genomic DNA. For example, the gene dose of the apolipoprotein E type 4 allele was correlated with the risk of Alzheimer's disease in late onset families (see, *e.g.*, Corder *et al.* (1993) *Science 261*: 921-923; mutation in blood coagulation factor V was associated with resistance to activated

10 protein C (see, *e.g.*, Bertina *et al.* (1994) *Nature 369*:64-67); resistance to HIV-1 infection has been shown in caucasian individuals bearing mutant alleles of the CCR-5 chemokine receptor gene (see, *e.g.*, Samson *et al.* (1996) *Nature 382*:722-725); and a hypermutable tract in antigen presenting cells (APC, such as macrophages), has been

15 identified in familial colorectal cancer in individuals of Ashkenzi jewish background (see, *e.g.*, Laken *et al.* (1997) *Nature Genet. 17*:79-83). There may be more than three million polymorphic sites in the human genome. Many have been identified, but not yet characterized or mapped or associated with a marker.

20 ### Single nucleotide polymorphisms (SNPs)

Much of the focus of genomics has been in the identification of SNPs, which are important for a variety of reasons. They allow indirect testing (association of haplotypes) and direct testing (functional variants). They are the most abundant and stable genetic markers. Common

25 diseases are best explained by common genetic alterations, and the natural variation in the human population aids in understanding disease, therapy and environmental interactions.

Currently, the only available method to identify SNPs in DNA is by sequencing, which is expensive, difficult and laborious. Furthermore,

30 once a SNP is discovered it must be validated to determine if it is a real

-3-

polymorphism and not a sequencing error. Also, discovered SNPs must then be evaluated to determine if they are associated with a particular phenotype. Thus, there is a need to develop new paradigms for identifying the genomic basis for disease and markers thereof. Therefore,

5   it is an object herein to provide methods for identifying the genomic basis of disease and markers thereof.

SUMMARY

Databases and methods using the databases are provided herein. The databases comprise sets of parameters associated with subjects in

10 populations selected only on the basis of being healthy (*i.e.*, where the subjects are mammals, such as humans, they are selected based upon apparent health and no detectable infections). The databases can be sorted based upon one or more of the selected parameters.

The databases are preferably relational databases, in which an

15 index that represents each subject serves to relate parameters, which are the data, such as age, ethnicity, sex, medical history, etc. and ultimately genotypic information, that was inputted into and stored in the database. The database can then be sorted according to these parameters. Initially, the parameter information is obtained from a questionnaire answered by

20 each subject from whom a body tissue or body fluid sample is obtained. As additional information about each sample is obtained, this information can be entered into the database and can serve as a sorting parameter.

The databases obtained from healthy individuals have numerous uses, such as correlating known polymorphisms with a phenotype or

25 disease. The databases can be used to identify alleles that are deleterious, that are beneficial, and that are correlated with diseases.

For purposes herein, genotypic information can be obtained by any method known to those of skill in the art, but is preferably obtained using mass spectrometry.

-4-

Also provided herein, is a new use for existing databases of subjects and genotypic and other parameters, such as age, ethnicity, race, and gender. Any database can be sorted according to the methods herein and alleles that exhibit statistically significant correlations with any

5   of the sorting parameters can be identified. It is noted, however, is noted, that the databases provided herein and randomly selected databases will perform better in these methods, since disease-based databases suffer numerous limitations, including their relatively small size, the homogeneity of the selected disease population, and the masking

10   effect of the polymorphism associated with the markers for which the database was selected. Hence, the healthy database provided herein, provides advantages not heretofore recognized or exploited. However, the methods provided herein can be used with a selected database, including disease-based databases, with or without sorting for the

15   discovery and correlation of polymorphisms. In addition, the databases provided herein represent a greater genetic diversity than the unselected databases typically utilized for the discovery of polymorphisms and thus allow for the enhanced discovery and correlation of polymorphisms.

The databases provided herein can be used for taking an identified

20   polymorphism, and ascertaining whether it changes in frequency when the data is sorted according to a selected parameter.

One use of these methods is correlating a selected marker with a particular parameter by following the occurrence of known genetic markers and then, having made this correlation, determining or identifying

25   correlations with diseases. Examples of this use are p53 and Lipoprotein Lipase polymorphism. As exemplified herein, known markers are shown

to have particular correlation with certain groups, such as a particular ethnicity or race or one sex. Such correlations will then permit development of better diagnostic tests and treatment regimens.

5   These methods are valuable for identifying one or more genetic markers whose frequency changes within the population as a function of age, ethnic group, sex or some other criteria. This can allow the identification of previously unknown polymorphisms and ultimately a gene or pathway involved in the onset and progression of disease.

10   The databases and methods provided herein permit, among other things, identification of components, particularly key components, of a disease process by understanding its genetic underpinnings and also permit an understanding of processes, such as individual drug responses. The databases and methods provided herein also can be used in methods involving elucidation of pathological pathways, in developing new

15 diagnostic assays, identifying new potential drug targets, and in identifying new drug candidates.

    The methods and databases can be used with experimental procedures, including, but are not limited to, *in silico* SNP identification, *in vitro* SNP identification/verification, genetic profiling of large populations,

20 and in biostatistical analyses and interpretations.

    Also provided herein, are combinations that contain a database provided herein and a biological sample from a subject in the database, and preferably biological samples from all subjects or a plurality of subjects in the database. Collections of the tissue and body fluid samples

25 are also provided.

    Also, provided herein, are methods for determining a genetic marker that correlates with age, comprising identifying a polymorphism and determining the frequency of the polymorphism with increasing age in a healthy population.

-6-

Further provided herein are methods for determining whether a genetic marker correlates with susceptibility to morbidity, early mortality, or morbidity and early mortality, comprising identifying a polymorphism and determining the frequency of the polymorphism with increasing age in

5   a healthy population.

Any of the methods herein described can be used out in a multiplex format.

Also provided are an apparatus and process for accurately identifying genetic information. It is another object of the herein that

10   genetic information be extracted from genetic data in a highly automated manner. Therefore, to overcome the deficiencies in the known conventional systems, a method and apparatus for identifying a biological sample is proposed.

Briefly, the method and system for identifying a biological sample

15   generates a data set indicative of the composition of the biological sample. In a particular example, the data set is DNA spectrometry data received from a mass spectrometer. The data set is denoised, and a baseline is deleted. Since possible compositions of the biological sample may be known, expected peak areas may be determined. Using the

20   expected peak areas, a residual baseline is generated to further correct the data set. Probable peaks are then identifiable in the corrected data set, which are used to identify the composition of the biological sample. In a disclosed example, statistical methods are employed to determine the probability that a probable peak is an actual peak, not an actual peak, or

25   that the data too inconclusive to call.

Advantageously, the method and system for identifying a biological sample accurately makes composition calls in a highly automated manner. In such a manner, complete SNP profile information, for example, may be collected efficiently. More importantly, the collected data is analyzed

30   with highly accurate results. For example, when a particular composition

-7-

is called, the result may be relied upon with great confidence. Such confidence is provided by the robust computational process employed .

DESCRIPTION OF THE DRAWINGS

Figure 1 depicts an exemplary sample bank. Panel 1 shows the samples as a function of sex and ethnicity. Panel 2 shows the caucasians as a function of age. Panel 3 shows the Hispanics as a function of age.

Figures 2A and 2C show an age- and sex-distribution of the 291S allele of the lipoprotein lipase gene in which a total of 436 males and 589 females were investigated. Figure 2B shows an age distribution for the 436 males.

Figure 3 is an exemplary questionnaire for population-based sample banking.

Figure 4 depicts processing and tracking of blood sample components.

Figure 5 depicts the allelic frequency of "sick" alleles and "healthy" alleles as a function of age. It is noted that the relative frequency of healthy alleles increases in a population with increasing age.

Figure 6 depicts the age-dependent distribution of ApoE genotypes (see, Schächter *et al.* (1994) *Nature Genetics 6*:29-32).

Figure 7A-D depicts age-related and genotype frequency of the p53 (tumor suppressor) codon 72 among the caucasian population in the database. *R72 and *P72 represent the frequency of the allele in the database population. R72, R72P, and P72 represent the genotypes of the individuals in the population. The frequency of the homozygous P72 allele drops from 6.7% to 3.7% with age.

Figure 8 depicts the allele and genotype frequencies of the p21 S31R allele as a function of age.

Figure 9 depicts the frequency of the FVII Allele 353Q in pooled versus individual samples.

-8-

Figure 10 depicts the frequency of the CETP (cholesterol ester transfer protein) allele in pooled versus individual samples

Figure 11 depicts the frequency of the plasminogen activator inhibitor-1 (PAI-1) 5G in pooled versus individual samples

5 Figure 12 shows mass spectra of the samples and the ethnic diversity of the PAI-1 alleles.

Figure 13 shows mass spectra of the samples and the ethnic diversity of the CETP 405 alleles.

Figure 14 shows mass spectra of the samples and the ethnic 10 diversity of the Factor VII 353 alleles.

Figure 15 shows ethnic diversity of PAI-1, CETP and Factor VII using the pooled DNA samples.

Figure 16 shows the p53-Rb pathway and the relationships among the various factors in the pathway.

15 Figure 17, which is a block diagram of a computer constructed to provide and process the databases described herein, depicts a typical computer system for storing and sorting the databases provided herein and practicing the methods provided herein.

Figure 18 is a flow diagram that illustrates the processing steps 20 performed using the computer illustrated in Figure 17, to maintain and provide access to the databases for identifying polymorphic genetic markers.

Figure 19 is a histogram showing the allele and genotype distribution in the age and sex stratified Caucasian population for the 25 AKAP10-1 locus. Bright green bars show frequencies in individuals younger than 40 years. Dark green bars show frequencies in individuals older than 60 years.

Figure 20 is a histogram showing the allele and genotype distribution in the age and sex stratified Caucasian population for the 30 AKAP10-5 locus. Bright green bars show frequencies in individuals

younger than 40 years; dark green bars show frequencies in individuals older than 60 years.

Figure 21 is a histogram showing the allele and genotype distribution in the age and sex stratified Caucasian population for the h-msrA locus. Genotype difference between male age groups is significant. Bright green bars show frequencies in individuals younger than 40 years. Dark green bars show frequencies in individuals older than 60 years.

Figure 22A-D is a sample data collection questionnaire used for the healthy database.

Figure 23 is a flowchart showing processing performed by the computing device of Figure 24 when performing genotyping of sense strands and antisense strands from assay fragments.

Figure 24 is a block diagram showing a system in accordance with the present invention;

Figure 25 is a flowchart of a method of identifying a biological sample in accordance with the present invention;

Figure 26 is a graphical representation of data from a mass spectrometer;

Figure 27 is a diagram of wavelet transformation of mass spectrometry data;

Figure 28 is a graphical representation of wavelet stage 0 hi data;

Figure 29 is a graphical representation of stage 0 noise profile;

Figure 30 is a graphical representation of generating stage noise standard deviations;

Figure 31 is a graphical representation of applying a threshold to data stages;

Figure 32 is a graphical representation of a sparse data set;

Figure 33 is a formula for signal shifting;

Figure 34 is a graphical representation of a wavelet transformation of a denoised and shifted signal;

-10-

Figure 35 is a graphical representation of a denoised and shifted signal;

Figure 36 is a graphical representation of removing peak sections;

Figure 37 is a graphical representation of generating a peak free signal ;

Figure 38 is a block diagram of a method of generating a baseline correction;

Figure 39 is a graphical representation of a baseline and signal;

Figure 40 is a graphical representation of a signal with baseline removed;

Figure 41 is a table showing compressed data;

Figure 42 is a flowchart of method for compressing data;

Figure 43 is a graphical representation of mass shifting;

Figure 44 is a graphical representation of determining peak width;

Figure 45 is a graphical representation of removing peaks;

Figure 46 is a graphical representation of a signal with peaks removed;

Figure 47 is a graphical representation of a residual baseline;

Figure 48 is a graphical representation of a signal with residual baseline removed;

Figure 49 is a graphical representation of determining peak height;

Figure 50 is a graphical representation of determining signal-to-noise for each peak;

Figure 51 is a graphical representation of determining a residual error for each peak;

Figure 52 is a graphical representation of peak probabilities;

Figure 53 is a graphical representation of applying an allelic ratio to peak probability;

Figure 54 is a graphical representation of determining peak probability

-11-

Figure 55 is a graphical representation of calling a genotype;

Figure 56 is a flowchart showing a statistical procedure for calling a genotype;

Figure 57 is a flowchart showing processing performed by the computing device of Figure 1 when performing standardless genotyping; and

Figure 58 is graphical representation of applying an allelic ratio to peak probability for standardless genotype processing.

DETAILED DESCRIPTION

Definitions

Unless defined otherwise, all technical and scientific terms used herein have the same meaning as is commonly understood by one of ordinary skill in the art to which this invention belongs. All patents, applications, published applications and other publications and sequences from GenBank and other databases referred to herein throughout the disclosure are incorporated by reference in their entirety.

As used herein, a biopolymer includes, but is not limited to, nucleic acid, proteins, polysaccharides, lipids and other macromolecules. Nucleic acids include DNA, RNA, and fragments thereof. Nucleic acids may be derived from genomic DNA, RNA, mitochondrial nucleic acid, chloroplast nucleic acid and other organelles with separate genetic material.

As used herein, morbidity refers to conditions, such as diseases or disorders, that compromise the health and well-being of an organism, such as an animal. Morbidity susceptibility or morbidity-associated genes are genes that, when altered, for example, by a variation in nucleotide sequence, facilitate the expression of a specific disease clinical phenotype. Thus, morbidity susceptibility genes have the potential, upon alteration, of increasing the likelihood or general risk that an organism will develop a specific disease.

As used herein, mortality refers to the statistical likelihood that an organism, particularly an animal, will not survive a full predicted lifespan. Hence, a trait or a marker, such as a polymorphism, associated with increased mortality is observed at a lower frequency in older than younger

5   segments of a population.

As used herein, a polymorphism, e.g. genetic variation, refers to a variation in the sequence of a gene in the genome amongst a population, such as allelic variations and other variations that arise or are observed. Thus, a polymorphism refers to the occurrence of two or more genetically

10   determined alternative sequences or alleles in a population. These differences can occur in coding and non-coding portions of the genome, and can be manifested or detected as differences in nucleic acid sequences, gene expression, including, for example transcription, processing, translation, transport, protein processing, trafficking, DNA

15   synthesis, expressed proteins, other gene products or products of biochemical pathways or in post-translational modifications and any other differences manifested amongst members of a population. A single nucleotide polymorphism (SNP) refers to a polymorphism that arises as the result of a single base change, such as an insertion, deletion or

20   change in a base.

A polymorphic marker or site is the locus at which divergence occurs. Such site may be as small as one base pair (an SNP). Polymorphic markers include, but are not limited to, restriction fragment length polymorphisms, variable number of tandem repeats (VNTR's),

25   hypervariable regions, minisatellites, dinucleotide repeats, trinucleotide repeats, tetranucleotide repeats and other repeating patterns, simple sequence repeats and insertional elements, such as Alu. Polymorphic forms also are manifested as different mendelian alleles for a gene. Polymorphisms may be observed by differences in proteins, protein

30   modifications, RNA expression modification, DNA and RNA methylation,

-13-

regulatory factors that alter gene expression and DNA replication, and any other manifestation of alterations in genomic nucleic acid or organelle nucleic acids.

As used herein, a healthy population, refers to a population of
5 organisms, including but are not limited to, animals, bacteria, viruses, parasites, plants, eubacteria, and others, that are disease free. The concept of disease-free is a function of the selected organism. For example, for mammals it refers to a subject not manifesting any disease state. Practically a healthy subject, when human, is defined as human
10 donor who passes blood bank criteria to donate blood for eventual use in the general population. These criteria are as follows: free of detectable viral, bacterial, mycoplasma, and parasitic infections; not anemic; and then further selected based upon a questionnaire regarding history (see Figure 3). Thus, a healthy population represents an unbiased population
15 of sufficient health to donate blood according to blood bank criteria, and not further selected for any disease state. Typically such individuals are not taking any medications. For plants, for example, it is a plant population that does not manifest diseases pathology associated with plants. For bacteria it is a bacterial population replicating without
20 environmental stress, such as selective agents, heat and other pathogens.

As used herein, a healthy database (or healthy patient database) refers to a database of profiles of subjects that have not been pre-selected for any particular disease. Hence, the subjects that serve as the source of data for the database are selected, according to predetermined
25 criteria, to be healthy. In contrast to other such databases that have been pre-selected for subjects with a particular disease or other characteristic, the subjects for the database provided herein are not so-selected. Also, if the subjects do manifest a disease or other condition, any polymorphism discovered or characterized should be related to an
30 independent disease or condition. In a preferred embodiment, where the

-14-

subjects are human, a healthy subject manifests no disease symptoms and meets criteria, such as those set by blood banks for blood donors.

Thus, the subjects for the database are a population of any organism, including, but are not limited to, animals, plants, bacteria, 5 viruses, parasites and any other organism or entity that has nucleic acid. Among preferred subjects are mammals, preferably, although not necessarily, humans. Such a database can capture the diversity of the a population, thus providing for discovery of rare polymorphisms.

As used herein, a profile refers to information relating to, but not 10 limited to and not necessarily including all of, age, sex, ethnicity, disease history, family history, phenotypic characteristics, such as height and weight and other relevant parameters. A sample collect information form is shown in Figure 22, which illustrates profile intent.

As used herein, a disease state is a condition or abnormality or 15 disorder that may be inherited or result from environmental stresses, such as toxins, bacterial, fungal and viral infections.

As used herein, set of non-selected subjects means that the subjects have not been pre-selected to share a common disease or other characteristic. They can be selected to be healthy as defined herein.

20 As used herein, a phenotype refers to a set of parameters that includes any distinguishable trait of an organism. A phenotype can be physical traits and can be, in instances in which the subject is an animal, a mental trait, such as emotional traits. Some phenotypes can be determined by observation elicited by questionnaires (see, *e.g.*, Figures 3 25 and 22) or by referring to prior medical and other records. For purposes herein, a phenotype is a parameter around which the database can be sorted.

As used herein, a parameter is any input data that will serve as a basis for sorting the database. These parameters will include phenotypic 30 traits, medical histories, family histories and any other such information

-15-

elicited from a subject or observed about the subject. A parameter may describe the subject, some historical or current environmental or social influence experienced by the subject, or a condition or environmental influence on someone related to the subject. Paramaters include, but are

5   not limited to, any of those described herein, and known to those of skill in the art.

As used herein, haplotype refers referes to two or polymorphism located on a single DNA strand. Hence, haplotyping refers to identification of two or more polymorphisms on a single DNA strand. Haplotypes can

10   be indicative of a phenotype. For some disorders a single polymorphism may suffice to indicate a trait; for others a plurality (i.e., a haplotype) may be needed. Haplotyping can be performed by isolating nucleic acid and separating the strands. In addition, when using enzymes such a certain nucleases, that produce, different size fragments from each strand, strand

15   separation is not needed for haplotyping.

As used herein, used herein, pattern with reference to a mass spectrum or mass spectrometric analyses, refers to a characteristic distribution and number of signals (such peaks or digital representations thereof).

20   As used herein, signal in the context of a mass spectrum and analysis thereof refers to the output data, which the number or relative number of moleucles having a particular mass. Signals include "peaks" and digital representations thereof.

As used herein, adaptor, when used with reference to haplotyping

25   use Fen ligase, refers to a nucleic acid that specifically hybridizes to a polymorphism of insterest. An adaptor can be partially double-stranded. An adaptor complex is formed when an adaptor hybridizes to its target.

As used herein, a target nucleic acid refers to any nucleic acid of interest in a sample. It can contain one or more nucleotides.

-16-

As used herein, standardless analysis refers to a determination based upon an internal standard. For example, the frequency of a polymorphism can be determined herein by comparing signals within a single mass spectrum.

5   As used herein, amplifying refers to means for increasing the amount of a bipolymer, especially nucleic acids. Based on the 5' and 3' primers that are chosen, amplication also serves to restrict and define the region of the genome which is subject to analysis. Amplification can be by any means known to those skilled in the art, including use of the

10   polymerase chain reaction (PCR) etc. Amplification, e.g., PCR must be done quantitatively when the frequency of polymorphism is required to be determined.

As used herein, cleaving refers to non-specific and specific fragmentation of a biopolymer.

15   As used herein, multiplexing refers to the simultaneous detection of more than one polymorphism. Methods for performing multiplexed reactions, particularly in conjunction with mass spectrometry are known (see, *e.g.*, U.S. Patent Nos. 6,043,031, 5,547,835 and International PCT application No. WO 97/37041).

20   As used herein, reference to mass spectrometry encompasss any suitable mass spectrometric format known to those of skill in the art. Such formats cinlude, but are not limited to, Matrix-Assisted Laser Desorption/Ionization, Time-of-Flight (MALDI-TOF), Electrospray (ES), IR-MALDI (see, *e.g.*, published International PCT application No.99/57318

25   and U.S. Patent No. 5,118,937), Ion Cyclotron Resonance (ICR), Fourier Transform and combinations thereof. MALDI, particular UV and IR, are among the preferred formats.

As used herein, mass spectrum refers to the presentation of data obtained from analyzing a biopolymer or fragment thereof by mass spectrometry either graphically or encoded numerically.

5   As used herein, a blood component is a component that is separated from blood and includes, but is not limited to red blood cells and platelets, blood clotting factors, plasma, enzymes, plasminogen, immunoglobulins. A cellular blood component is a component of blood, such as a red blood cell, that is a cell. A blood protein is a protein that is normally found in blood. Examples of such proteins are blood factors VII

10 and VIII. Such proteins and components are well-known to those of skill in the art.

As used herein, plasma can be prepared by any method known to those of skill in the art. For example, it can be prepared by centrifuging blood at a force that pellets the red cells and forms an interface between

15 the red cells and the buffy coat, which contains leukocytes, above which is the plasma. For example, typical platelet concentrates contain at least about 10% plasma.

Blood may be separated into its components, including, but not limited to, plasma, platelets and red blood cells by any method known to

20 those of skill in the art. For example, blood can be centrifuged for a sufficient time and at a sufficient acceleration to form a pellet containing the red blood cells. Leukocytes collect primarily at the interface of the pellet and supernatant in the buffy coat region. The supernatant, which contains plasma, platelets, and other blood components, may then be

25 removed and centrifuged at a higher acceleration, whereby the platelets pellet.

-18-

As used herein, p53 is a cell cycle control protein that assesses DNA damage and acts as a transcription factor regulation gene which control cell growth, DNA repair and apoptosis. The p53 mutations have been found in a wide variety of different cancers, including all of the
5   different types of leukemia, with varying frequency. The loss of normal p53 functions results in genomic instability and uncontrolled growth of the host cell.

As used herein, p21 is a cyclin-dependent kinase inhibitor, associated with G1 phase arrest of normal cells. Expression triggers
10   apoptosis or programmed cell death and has been associated with Wilms' tumor, a pediatric kidney cancer.

As used herein, Factor VII is a serine protease involved the extrinsic blood coagulation cascade. This factor is activated by thrombin and works with tissue factor (Factor III) in the processing of Factor X to
15   Factor Xa. Evidence has supported an association between polymorphisms in the gene and increase Factor VII activity which can result in an elevated risk of ischemic cardiovascular disease including myocardial infarction.

As used herein, a relational database stores information in a form
20   representative of matrices, such as two-dimensional tables, including rows and columns of data, or higher dimensional matrices. For example, in one embodiment, the relational database has separate tables each with a parameter. The tables are linked with a record number, which also acts as an index. The database can be searched or sorted by using data in the
25   tables and is stored in any suitable storage medium, such as floppy disk, CD rom disk, hard drive or other suitable medium.

As used herein, a bar codes refers any array of optically readable marks of any desired size and shape that are arranged in a reference context or frame of, preferably, although not necessarily, one or more
30   columns and one or more rows. For purposes herein, the bar code refers

to any symbology, not necessary "bar" but may include dots, characters or any symbol or symbols.

As used herein, symbology refers to an identifier code or symbol, such as a bar code, that is linked to a sample. The index will reference
5 each such symbology. The symbology is any code known or designed by the user. The symbols are associated with information stored in the database. For example, each sample can be uniquely identified with an encoded symbology. The parameters, such as the answers to the questions and subsequent genotypic and other information obtained upon
10 analysis of the samples is included in the database and associated with the symbology. The database is stored on any suitable recording medium, such as a hard drive, a floppy disk, a tape, a CD ROM, a DVD disk and any other suitable medium.

DATABASES
15 Human genotyping is currently dependent on collaborations with hospitals, tissues banks and research institutions that provide samples of disease tissue. This approach is based on the concept that the onset and/or progression of diseases can be correlated with the presence of a polymorphisms or other genetic markers. This approach does not
20 consider that disease correlated with the presence of specific markers and the absence of specific markers. It is shown herein that identification and scoring of the appearance and disappearance of markers is possible only if these markers are measured in the background of healthy subjects where the onset of disease does not mask the change in polymorphism
25 occurrence. Databases of information from disease populations suffer from small sample size, selection bias and heterogeneity. The databases provided herein from healthy populations solve these problems by permitting large sample bands, simple selection methods and diluted heterogeneity.

Provided herein are first databases of parameters, associated with non-selected, particularly healthy, subjects. Also provided are combinations of the databases with indexed samples obtained from each of the subjects. Further provided are databases produced from the first

5   databases. These contain in addition to the original parameters information, such as genotypic information, including, but are not limited to, genomic sequence information, derived from the samples.

The databases, which are herein designated healthy databases, are so-designated because they are not obtained from subjects pre-selected

10 for a particular disease. Hence, although individual members may have a disease, the collection of individuals is not selected to have a particular disease.

The subjects from whom the parameters are obtained comprise either a set of subjects who are randomly selected across, preferably, all

15 populations, or are pre-selected to be disease-free or healthy. As a result, the database is not selected to be representative of any pre-selected phenotype, genotype, disease or other characteristic. Typically the number of subjects from which the database is prepared is selected to produce statistically significant results when used in the methods

20 provided herein. Preferably, the number of subjects will be greater than 100, more preferably greater than 200, yet more preferably greater than 1000. The precise number can be empirically determined based upon the frequency of the parameter(s) that be used to sort the database. Generally the population can have at least 50, at least 100, at least 200,

25 at least 500, at least 1000, at least 5000 or at least 10,000 or more subjects.

Upon identification of a collection of subjects, information about each subject is recorded and associated with each subject as a database. The information associated with each of the subjects, includes, but is not

30 limited to, information related to historical characteristics of the subjects,

-21-

phenotypic characteristics and also genotypic characteristics, medical characteristics and any other traits and characteristics about the subject that can be determined. This information will serve as the basis for sorting the database.

5   In an exemplary embodiment, the subjects are mammals, such as humans, and the information relates to one or more of parameters, such as age, sex, medical history, ethnicity and any other factor. Such information, when the animals are humans, for example, can be obtained by a questionnaire, and by observations about the individual, such as hair

10   color, eye color and other characteristics. Genotypic information will be obtained from tissue or other body and body fluid samples from the subject.

The healthy genomic database can include profiles and polymorphisms from healthy individuals from a library of blood samples

15   where each sample in the library is an individual and separate blood or other tissue sample. Each sample in the database is profiled as to the sex, age, ethnic group, and disease history of the donor.

The databases are generated by first identifying healthy populations of subjects and obtaining information about each subject that will serve

20   as the sorting parameters for the database. This information is preferably entered into a storage medium, such as the memory of a computer.

The information obtained about each subject in a population used for generating the database is stored in a computer memory or other suitable storage medium. The information is linked to an identifier

25   associated with each subject. Hence the database will identify a subject, for example by a datapoint representative of a bar code, and then all information, such as the information from a questionnaire, regarding the individual is associated with the datapoint. As the information is collected the database is generated.

-22-

Thus, for example, profile information, such as subject histories obtained from questionnaires, is collected in the database. The resulting database can be sorted as desired, using standard software, such as by age, sex and/or ethnicity. An exemplary questionnaire for subjects from

5 whom samples are to be obtained is shown in Figures 22A-D. Each questionnaire preferably is identified by a bar code, particularly a machine readable bar code for entry into the database. After a subject provides data and is deemed to be healthy (*i.e.*, meets standards for blood donation), the data in the questionnaire is entered into the database and

10 is associated with the bar code. A tissue, cell or blood sample is obtained from the subject.

Figure 4 exemplifies processing and tracking of blood sample components. Each component is tracked with a bar code, dated, is entered into the database and associated with the subject and the profile

15 of the subject. Typically, the whole blood is centrifuged to produce plasma, red blood cells (which pellet) and leukocytes found in the buffy coat which layers in between. Various samples are obtained and coded with a bar code and stored for use as needed.

Samples are collected from the subjects. The samples include, but

20 are not limited to, tissues, cells, and fluids, such as nucleic acid, blood, plasma, amniotic fluid, synovial fluid, urine, saliva, aqueous humor, sweat, sperm samples and cerebral spinal fluid. It is understood that the particular set of samples depends upon the organisms in the population.

Once samples are obtained the collection can be stored and, in

25 preferred embodiments, each sample is indexed with an identifier, particularly a machine readable code, such as a bar code. For analyses, the samples or components of the samples, particularly biopolymers and small molecules, such as nucleic acids and/or proteins and metabolites, are isolated.

After samples are analyzed, this information is entered into the database in the memory of the storage medium and associated with each subject. This information includes, but is not limited to, genotypic information. Particularly, nucleic acid sequence information and other
5   information indicative of polymorphisms, such as masses of PCR fragments, peptide fragment sequences or masses, spectra of biopolymers and small molecules and other indicia of the structure or function of a gene, gene product or other marker from which the existence of a polymorphism within the population can be inferred.
10   In an exemplary embodiment, a database can be derived from a collection of blood samples. For example, Figure 1 (see, also Figure 10) shows the status of a collection of over 5000 individual samples. The samples were processed in the laboratory following SOP (standard operating procedure) guidelines. Any standard blood processing protocol
15   may be used.

For the exemplary database described herein, the following criteria were used to select subjects:

No testing is done for infectious agents.

Age: At least 17 years old

20   Weight: Minimum of 110 pounds

Permanently Disqualified:

History of hepatitis (after age 11)

Leukemia Lymphoma

Human immunodeficiency virus (HIV),AIDS

25   Chronic kidney disease

Temporarily Disqualified:

Pregnancy - until six weeks after delivery, miscarriage or abortion

Major surgery or transfusions - for one year

Mononucleosis - until complete recovery

30   Prior whole blood donation - for eight weeks

-24-

Antibiotics by injection for one week; by mouth, for forty-eight

hours, except antibiotics for skin complexion;

5 year Deferment:

Internal cancer and skin cancer if it has been removed, is healed

5   and there is no recurrence

These correspond to blood bank criteria for donating blood and represent

a healthy population as defined herein for a human healthy database.

Structure of the database

Any suitable database structure and format known to those of skill

10   in the art may be employed. For example, a relational database is a

preferred format in which data is stored as matrices or tables of the

parameters linked by an indexer that identifies each subject. Software for

preparing and manipulating, including sorting the database, can be readily

developed or adapted from commercially available software, such as

15   Microsoft Access.

Quality control

Quality control procedures can be implemented. For example, after

collection of samples, the quality of the collection in the bank can be

assessed. For example, mix-up of samples can be checked by testing for

20   known markers, such as sex. After samples are separated by ethnicity,

samples are randomly tested for a marker associated with a particular

ethnicity, such as HLA DQA1 group specific component, to assess

whether the samples have been properly sorted by ethnic group. An

exemplary sample bank is depicted in Figure 4.

25   **Obtaining genotypic data and other parameters for the database**

After informational and historical parameters are entered into the

database, material from samples obtained from each subject, is analyzed.

Analyzed material include proteins, metabolites, nucleic acids, lipids and

any other desired constituent of the material. For example, nucleic acids,

30   such as genomic DNA, can be analyzed by sequencing.

-25-

Sequencing can be performed using any method known to those of skill in the art. For example, if a polymorphism is identified or known, and it is desired to assess its frequency or presence among the subjects in the database, the region of interest from each sample can be isolated,
5 such as by PCR or restriction fragments, hybridization or other suitable method known to those of skill in the art and sequenced. For purposes herein, sequencing analysis is preferably effected using mass spectrometry (see, *e.g.*, U.S. Patent Nos. 5,547,835, 5,622,824, 5,851,765, and 5,928,906). Nucleic acids can also be sequence by
10 hybridization (see, *e.g.*, U.S. Patent Nos. 5,503,980, 5,631,134, 5,795,714) and including analysis by mass spectrometry (see, U.S. application Serial Nos. 08/419,994 and 09/395,409).

In other detection methods, it is necessary to first amplify prior to identifying the allelic variant. Amplification can be performed, e.g., by
15 PCR and/or LCR, according to methods known in the art. In one embodiment, genomic DNA of a cell is exposed to two PCR primers and amplification for a number of cycles sufficient to produce the required amount of amplified DNA. In preferred embodiments, the primers are located between 150 and 350 base pairs apart.

20 Alternative amplification methods include: self sustained sequence replication (Guatelli, J. C. et al., 1990, Proc. Natl. Acad. Sci. U.S.A. 87:1874-1878), transcriptional amplification system (Kwoh, D. Y. et al., 1989, Proc. Natl. Acad. Sci. U.S.A. 86:1173-1177), Q-Beta Replicase (Lizardi, P. M. et al., 1988, Bio/Technology 6:1197), or any other nucleic
25 acid amplification method, followed by the detection of the amplified molecules using techniques well known to those of skill in the art. These detection schemes are especially useful for the detection of nucleic acid molecules if such molecules are present in very low numbers.

Nucleic acids can also be analyzed by detection methods and
30 protocols, particularly those that rely on mass spectrometry (see, *e.g.*,

-26-

U.S. Patent No. 5,605,798, 6,043,031, allowed copending U.S.
application Serial No. 08/744,481, U.S. application Serial No.
08/990,851 and International PCT application No. WO 99/31273,
International PCT application No. WO 98/20019). These methods can be
5   automated (see, *e.g.*, copending U.S. application Serial No. 09/285,481
and published International PCT application No. PCT/US00/08111, which
describes an automated process line). Preferred among the methods of
analysis herein are those involving the primer oligo base extension
(PROBE) reaction with mass spectrometry for detection (described herein
10 and elsewhere, see *e.g.*, U.S. Patent No. 6,043,031; see, also U.S.
application Serial Nos. 09/287,681, 09/287,682, 09/287,141 and
09/287,679, allowed copending U.S. application Serial No. 08/744,481,
International PCT application No. PCT/US97/20444, published as
International PCT application No. WO 98/20019, and based upon U.S.
15 application Serial Nos. 08/744,481, 08/744,590, 08/746,036,
08/746,055, 08/786,988, 08/787,639, 08/933,792, 08/746,055,
08/786,988 and 08/787,639; see, also U.S. application Serial No.
09/074,936, U.S. Patent No. 6,024,925, and U.S. application Serial Nos.
08/746,055 and 08/786,988, and published International PCT application
20 No. WO 98/20020)

     A preferred format for performing the analyses is a chip based
format in which the biopolymer is linked to a solid support, such as a
silicon or silicon-coated substrate, preferably in the form of an array.
More preferably, when analyses are performed using mass spectrometry,
25 particularly MALDI, small nanoliter volumes of sample are loaded on, such
that the resulting spot is about, or smaller than, the size of the laser spot.
It has been found that when this is achieved, the results from the mass
spectrometric analysis are quantitative. The area under the signals in the
resulting mass spectra are proportional to concentration (when normalized
30 and corrected for background). Methods for preparing and using such

-27-

chips are described in U.S. Patent No. 6,024,925, co-pending U.S. application Serial Nos. 08/786,988, 09/364,774, 09/371,150 and 09/297,575; see, also U.S. application Serial No. PCT/US97/20195, which published as WO 98/20020. Chips and kits for performing these

5 analyses are commercially available from SEQUENOM under the trademark MassARRAY. MassArray relies on the fidelity of the enzymatic primer extension reactions combined with the miniaturized array and MALDI-TOF (Matrix-Assisted Laser Desorption Ionization-Time of Flight) mass spectrometry to deliver results rapidly. It accurately distinguishes

10 single base changes in the size of DNA fragments associated with genetic variants without tags.

The methods provided herein permit quantitative determination of alleles. The areas under the signals in the mass spectra can be used for quantitative determinations. The frequency is determined from the ratio

15 of the signal to the total area of all of the spectrum and corrected for background. This is possible because of the PROBE technology as described in the above applications incorporated by reference herein.

Additional methods of analyzing nucleic acids include amplification-based methods including polymerase chain reaction (PCR), ligase chain

20 reaction (LCR), mini-PCR, rolling circle amplification, autocatalytic methods, such as those using Qβ replicase, TAS, 3SR, and any other suitable method known to those of skill in the art.

Other methods for analysis and identification and detection of polymorphisms, include but are not limited to, allele specific probes,

25 Southern analyses, and other such analyses.

The methods described below provide ways to fragment given amplified or non-amplified nucleotide sequences thereby producing a set of mass signals when mass spectrometry is used to analyze the fragment mixtures.

Amplified fragments are yielded by standard polymerase chain methods (US 4,683,195 and 4,683,202). The fragmentation method involves the use of enzymes that cleave single or double strands of DNA and enzymes that ligate DNA. The cleavage enzymes can be glycosylases, nickases,

5   and site-specific and non site-specific nucleases with the most preferred enzymes being glycosylases, nickases, and site-specific nucleases.

Glycosylase Fragmentation Method

DNA glycosylases specifically remove a certain type of nucleobase from a given DNA fragment. These enzymes can thereby produce abasic

10   sites, which can be recognized either by another cleavage enzyme, cleaving the exposed phosphate backbone specifically at the abasic site and producing a set of nucleobase specific fragments indicative of the sequence, or by chemical means, such as alkaline solutions and or heat. The use of one combination of a DNA glycosylase and its targeted

15   nucleotide would be sufficient to generate a base specific signature pattern of any given target region.

Numerous DNA glcosylases are known, For example, a DNA glycosylase can be uracil-DNA glycolsylase (UDG) , 3-methyladenine DNA glycosylase, 3-methyladenine DNA glycosylase II, pyrimidine hydrate-DNA

20   glycosylase, FaPy-DNA glycosylase, thymine mismatch-DNA glycosylase, hypoxanthine-DNA glycosylase, 5-Hydroxymethyluracil DNA glycosylase (HmUDG), 5-Hydroxymethylcytosine DNA glycosylase, or 1,N6-etheno-adenine DNA glycosylase (see, *e.g.*,, U.S. Patent Nos. 5,536,649, 5,888, 795, 5,952,176 and 6,099,553, International PCT application Nos. WO

25   97/03210, WO 99/54501; see, also, Eftedal et al. (1993) Nucleic Acids Res 21:2095-2101, Bjelland and Seeberg (1987) Nucleic Acids Res. 15:2787-2801, Saparbaev et al. (1995) Nucleic Acids Res. 23:3750-3755, Bessho (1999) Nucleic Acids Res. 27:979-983) corresponding to the enzyme's modified nucleotide or nucleotide analog target. A preferred

30   glycosylase is uracil-DNA glycolsylase (UDG).

-29-

Uracil, for example, can be incorporated into an amplified DNA molecule by amplifying the DNA in the presence of normal DNA precursor nucleotides (e.g. dCTP, dATP, and dGTP) and dUTP. When the amplified product is treated with UDG, uracil residues are cleaved. Subsequent
5   chemical treatment of the products from the UDG reaction results in the cleavage of the phosphate backbone and the generation of nucleobase specific fragments. Moreover, the separation of the complementary strands of the amplified product prior to glycosylase treatment allows complementary patterns of fragmentation to be generated. Thus, the use
10   of dUTP and Uracil DNA glycosylase allows the generation of T specific fragments for the complementary strands, thus providing information on the T as well as the A positions within a given sequence. Similar to this, a C-specific reaction on both (complementary) strands (i.e. with a C-specific glycosylase) yields information on C as well as G positions within
15   a given sequence if the fragmentation patterns of both amplification strands are analyzed separately. Thus, with the glycosylase method and mass spectrometry, a full series of A, C, G and T specific fragmentation patterns can be analyzed.

Nickase Fragmentation Method

20   A DNA nickase, or DNase, can be used recognize and cleave one strand of a DNA duplex. Numerous nickases are known. Among these, for example, are nickase NY2A nickase and NYS1 nickase (Megabase) with the following cleavage sites:

NY2A: 5'...R AG...3'

25   3'...Y TC...5' where R = A or G and Y = C or T

NYS1: 5'... CC[A/G/T]...3'

              3'... GG[T/C/A]...5'.

Fen-Ligase Fragmentation Method

The Fen-ligase method involves two enzymes: Fen-1 enzyme and a
30   ligase.

The Fen-1 enzyme is a site-specific nuclease known as a "flap" endonuclease (US 5,843,669, 5,874,283, and 6,090,606). This enzymes recognizes and cleaves DNA "flaps" created by the overlap of two oligonucleotides hybridized to a target DNA strand. This cleavage is

5   highly specific and can recognize single base pair mutations, permitting detection of a single homologue from an individual heterozygous at one SNP of interest and then genotyping that homologue at other SNPs occurring within the fragment. Fen-1 enzymes can be Fen-1 like nucleases e.g. human, murine, and Xenopus XPG enzymes and yeast RAD2

10 nucleases or Fen-1 endonucleases from, for example, *M. jannaschii*, *P. furiosus*, and *P. woesei*. Among preferred enzymes are the Fen-1 enzymes.

The ligase enzyme forms a phosphodiester bond between two double stranded nucleic acid fragments. The ligase can be DNA Ligase I

15 or DNA Ligase III (see, *e.g.*, U.S. Patent Nos. US 5,506,137, 5,700,672, 5,858,705 and 5,976,806; see, also, Waga, *et al.* (1994) J. Biol. Chem. 269:10923-10934, Li et al. (1994) Nucleic Acids Res. 22:632-638, Arrand et al. (1986) J. Biol. Chem. 261:9079-9082, Lehman (1974) Science 186:790-797, Higgins and Cozzarelli (1979) Methods Enzymol.

20 68:50-71, Lasko et al. (1990) Mutation Res. 236:277-287, and Lindahl and Barnes (1992) Ann. Rev. Biochem. 61:251-281). Thermostable ligase (Epicenter Technologies), where "thermostable" denotes that the ligase retains activity even after exposure to temperatures necessary to separate two strands of DNA, are among preferred ligases for use herein.

25 **Type IIS Enzyme Fragmentation Method**

Restriction enzymes bind specifically to and cleave double-stranded DNA at specific sites within or adjacent to a particular recognition sequence. These enzymes have been classified into three groups (e.g. Types I, II, and III) as known to those of skill in the art. Because of the

30 properties of type I and type III enzymes, they have not been widely used

-31-

in molecular biological applications. Thus, for this invention type II enzymes are preferred. Of the thousands of restriction enzymes known in the arts, there are 179 different type II specificities. Of the 179 unique type II restriction endonucleases, 31 have a 4-base recognition sequence,

5   11 have a 5-base recognition sequence, 127 have a 6-base recognition sequence, and 10 have recognition sequences of greater than six bases (US 5,604,098). Of category type II enzymes, type IIS is preferred.

Type IIS enzymes can be *Alw* XI, *Bbv* I, *Bce* 83, *Bpm* I, *Bsg* I, *Bsm* AI, *Bsm* FI, *Bsa* I, *Bcc* I, *Bcg* I, *Ear* I, *Eco* 57I, *Esp* 3I, *Fau* I, *Fok* I, *Gsu* I,

10   *Hga* I, *Mme* I, *Mbo* II, *Sap* I, and the like. The preferred type IIS enzyme is Fok I.

The Fok I enzyme endonuclease is an exemplary well characterized member of the Type IIS class (see, *e.g.*, U.S. Patent Nos. 5,714,330, 5,604,098, 5,436,150, 6,054,276 and 5,871,911; see, also, Szybalski

15   et al. (1991) Gene 100:13-26, Wilson and Murray (1991) Ann. Rev. Genet. 25:585-627, Sugisaki et al. (1981) Gene 16:73-78, Podhajska and Szalski (1985) Gene 40:175-182. Fok I recognizes the sequence 5'GGATG-3' and cleaves DNA accordingly. Type IIS restriction sites can be introduced into DNA targets by incorporating the site into primers used

20   to amplify such targets. Fragments produced by digestion with Fok I are site specific and can be analyzed by mass spectrometry methods such as MALDI-TOF mass spectrometry, ESI-TOF mass spectrometry, and any other type of mass spectrometry well known to those of skill in the art.

Once a polymorphism has been found to correlatate with a

25   parameter such as age. The possibility of false results due to allelic dropout is examined by doing comparative PCR in an adjacent region of the genome.

Analyses

In using the database, allelic frequencies can be determined across the population by analyzing each sample in the population individually, determining the presence or absence of allele or marker of interest in each
5   individual sample, and then determining the frequency of the marker in the population. The database can then be sorted (stratified) to identify any correlations between the allele and a selected parameter using standard statistical analysis. If a correlation is observed, such as a decrease in a particular marker with age or correlation with sex or other
10 parameter, then the marker is a candidate for further study, such as genetic mapping to identify a gene or pathway in which it is involved. The marker may then be correlated, for example, with a disease. Haplotying can also be carried out. Genetic mapping can be effected using standard methods and may also require use of databases of others,
15 such as databases previously determined to be associated with a disorder.

Exemplary analyses have been performed and these are shown in the figures, and discussed herein.

Sample pooling

20 It has been found that using the databases provided herein, or any other database of such information, substantially the same frequencies that were obtained by examining each sample separately can be obtained by pooling samples, such as in batches of 10, 20, 50, 100, 200, 500, 1000 or any other number. A precise number may be determined
25 empirically if necessary, and can be as low as 3.

In one embodiment, the frequency of genotypic and other markers can be obtained by pooling samples. To do this a target population and a genetic variation to be assessed is selected, a plurality of samples of biopolymers are obtained from members of the population, and the
30 biopolymer from which the marker or genotype can be inferred is

-33-

determined or detected. A comparison of samples tested in pools and individually and the sorted results therefrom are shown in Figure 9, which shows frequency of the factor VII Allele 353Q. Figure 10 depicts the frequency of the CETP Allele CETP in pooled versus individual samples.

5   Figure 15 shows ethnic diversity among various ethnic groups in the database using pooled DNA samples to obtain the data. Figures 12-14 show mass spectra for these samples.

Pooling of test samples has application not only to the healthy databases provided herein, but also to use in gathering data for entry into

10   any database of subjects and genotypic information, including typical databases derived from diseased populations. What is demonstrated herein, is the finding that the results achieved are statistically the same as the results that would be achieved if each sample is analyzed separately. Analysis of pooled samples by a method, such as the mass spectrometric

15   methods provided herein, permits resolution of such data and quantitation of the results.

For factor VII the R53Q acid polymorphism was assessed. In Figure 9, the "individual" data represent allelic frequency observed in 92 individuals reactions. The pooled data represent the allelic frequency of

20   the same 92 individuals pooled into a single probe reaction. The concentration of DNA in the samples of individual donors is 250 nanograms. The total concentration of DNA in the pooled samples is also 250 nanograms, where the concentration of any individual DNA is 2.7 nanograms.

25   It also was shown that it is possible to reduce the DNA concentration of individuals in a pooled samples from 2.7 nanograms to 0.27 nanograms without any change in the quality of the spectrum or the ability to quantitate the amount of sample detected. Hence low concentrations of sample may be used in the pooling methods.

Use of the databases and markers identified thereby

The successful use of genomics requires a scientific hypothesis (*i.e.*, common genetic variation, such as a SNP), a study design (*i.e.*, complex disorders), samples and technology, such as the chip-based

5   mass spectrometric analyses (see, *e.g.*, U.S. Patent No. 5,605,798, U.S. Patent No. 5,777,324, U.S. Patent No. 6,043,031, allowed copending U.S. application Serial No. 08/744,481, U.S. application Serial No. 08/990,851, International PCT application No. WO 98/20019, copending U.S. application Serial No. 09/285,481, which describes an automated

10 process line for analyses; see, also, U.S. application Serial Nos. 08/617,256, 09/287,681, 09/287,682, 09/287,141 and 09/287,679, allowed copending U.S. application Serial No. 08/744,481, International PCT application No. PCT/US97/20444, published as International PCT application No. WO 98/20019, and based upon U.S. application Serial

15 Nos. 08/744,481, 08/744,590, 08/746,036, 08/746,055, 08/786,988, 08/787,639, 08/933,792, 08/746,055, 09/266,409, 08/786,988 and 08/787,639; see, also U.S. application Serial No. 09/074,936). All of these aspects can be used in conjunction with the databases provided herein and samples in the collection.

20 The databases and markers identified thereby can be used, for example, for identification of previously unidentified or unknown genetic markers and to identify new uses for known markers. As markers are identified, these may be entered into the database to use as sorting parameters from which additional correlations may be determined.

25 **Previously unidentified or unknown genetic markers**

The samples in the healthy databases can be used to identify new polymorphisms and genetic markers, using any mapping, sequencing, amplification and other methodologies, and in looking for polymorphisms among the population in the database. The thus-identified polymorphism

30 can then be entered into the database for each sample, and the database

sorted (stratified) using that polymorphism as a sorting parameter to identify any patterns and correlations that emerge, such as age correlated changes in the frequency of the identified marker. If a correlation is identified, the locus of the marker can be mapped and its function or effect assessed or deduced.

5

Thus, the databases here provide means for:

identification of significantly different allelic frequencies of genetic factors by comparing the occurrence or disappearance of the markers with increasing age in population and then associating the markers with a disease or a biochemical pathway;

10

identification of significantly different allelic frequencies of disease causing genetic factors by comparing the male with the female population or comparing other selected stratified populations and associating the markers with a disease or a biochemical pathway;

15

identification of significantly different allelic frequencies of disease causing genetic factors by comparing different ethnic groups and associating the markers with a disease or a biochemical pathway that is known to occur in high frequency in the ethnic group;

profiling potentially functional variants of genes through the general panmixed population stratified according to age, sex, and ethnic origin and thereby demonstrating the contribution of the variant genes to the physical condition of the investigated population;

20

identification of functionally relevant gene variants by gene disequilibrium analysis performed within the general panmixed population stratified according to age, sex, and ethnic origin and thereby demonstrating their contribution to the physical condition of investigated population;

25

identification of potentially functional variants of chromosomes or parts of chromosomes by linkage disequilibrium analysis performed within the general panmixed population stratified according to age, sex, and

30

-36-

ethnic origin and thereby demonstrating their contribution to the physical condition of investigated population.

Uses of the identified markers and known markers

The databases may also be used in conjunction with known markers and sorted to identify any correlations. For example, the databases can be used for:

determination and evaluation of the penetrance of medically relevant polymorphic markers;

determination and evaluation of the diagnostic specificity of medically relevant genetic factors;

determination and evaluation of the positive predictive value of medically relevant genetic factors;

determination and evaluation of the onset of complex diseases, such as, but are not limited to, diabetes, hypertension, autoimmune diseases, arteriosclerosis, cancer and other diseases within the general population with respect to their causative genetic factors;

delineation of the appropriate strategies for preventive disease treatment;

delineation of appropriate timelines for primary disease intervention;

validation of medically relevant genetic factors identified in isolated populations regarding their general applicability;

validation of disease pathways including all potential target structures identified in isolated populations regarding their general applicability; and

validation of appropriate drug targets identified in isolated populations regarding their general applicability.

Among the diseases and disorders for which polymorphisms may be linked include, those linked to inborn errors of metabolism, acquired metabolic disorders, intermediary metabolism, oncogenesis pathways, blood clotting pathways, and DNA synthetic and repair pathways DNA

-37-

repair/replication/transcription factors and activities, *e.g.*, such as genes related to oncogenesis, aging and genes involved in blood clotting and the related biochemical pathways that are related to thrombosis, embolism, stroke, myocardial infarction, angiogenesis and oncogenesis.

5   For example, a number of diseases are caused by or involve deficient or defective enzymes in intermediary metabolism (see, <u>e.g.</u>, Tables 1 and 2, below) that result, upon ingestion of the enzyme substrates, in accumulation of harmful metabolites that damage organs and tissues, particularly an infant's developing brain and other organs,

10   resulting in mental retardation and other developmental disorders. **Identification of markers and genes for such disorders is of great interest.**

Model systems

Several gene systems, p21, p53 and Lipoprotein Lipase polymorphism (N291S), were selected. The p53 gene is a tumor suppressor gene that is mutated in diverse tumor types. One common

15   allelic variant occurs at codon 72. A polymorphism that has been identified in the p53 gene, i.e., the R72P allele, results in an amino acid exchange, arginine to proline, at codon 72 of the gene.

Using diseased populations, it has been shown that there are ethnic

20   differences in the allelic distribution of these alleles among African-Americans and Caucasians in the U.S. The results here support this finding and also demonstrate that the results obtained with a healthy database are meaningful (see, Figure 7B).

The 291S allele leads to reduced levels of high density lipoprotein

25   cholesterol (HDL-C) that is associated with an increased risk of males for arteriosclerosis and in particular myocardial infarction (see, Reymer *et al.* (1995) *Nature Genetics 10*:28-34).

Both genetic polymorphisms were profiled within a part of the Caucasian population-based sample bank. For the polymorphism located

30   in the lipoprotein lipase gene a total of 1025 unselected individuals (436

-38-

males and 589 females) were tested. Genomic DNA was isolated from blood samples obtained from the individuals.

As shown in the Examples and figures, an exemplary database containing about 5000 subjects, answers to the questionnaire (see Figure 3), and genotypic information has been stratified. A particular known allele has been selected, and the samples tested for the marker using mass spectrometric analyses, particularly PROBE (see the EXAMPLES) to identify polymorphisms in each sample. The population in the database has been sorted according to various parameters and correlations have been observed. For example, FIGURES 2A-C, show sorting of the data by age and sex for the Lipoprotein Lipase gene in the Caucasian population in the database. The results show a decrease in the frequency of the allele with age in males but no such decrease in females. Other alleles that have been tested against the database, include, alleles of p53, p21 and factor VII. Results when sorted by age are shown in the figures.

These examples demonstrate an effect of altered frequency of disease causing genetic factors within the general population. The scientific interpretation of those results allows prediction of medical relevance of polymorphic genetic alterations. In addition, conclusions can be drawn with regard to their penetrance, diagnostic specificity, positive predictive value, onset of disease, most appropriate onset of preventive strategies, and the general applicability of genetic alterations identified in isolated populations to panmixed populations.

Therefore, an age- and sex-stratified population-based sample bank that is ethnically homogenous is a suitable tool for rapid identification and validation of genetic factors regarding their potential medical utility.

Exemplary computer system for creating, storing and processing the databases

Systems

Systems, including computers, containing the databases are provided herein. The computers and databases can be used in
5   conjunction, for example, with the APL system (see, copending U.S. application Serial No. 09/285,481), which is an automated system for analyzing biopolymers, particularly nucleic acids. Results from the APL system can be entered into the database.

Any suitable computer system may be used. The computer system
10   may be integrated into systems for sample analysis, such as the automated process line described herein (see, *e.g.*, copending U.S. application Serial No. 09/285,481).

Figure 17 is a block diagram of a computer constructed in to provide and process the databases described herein. The processing that
15   maintains the database and performs the methods and procedures may be performed on multiple computers all having a similar construction, or may be performed by a single, integrated computer. For example, the computer through which data is added to the database may be separate from the computer through which the database is sorted, or may be
20   integrated with it. In either arrangement, the computers performing the processing may have a construction as illustrated in Figure 17.

Figure 17 is a block diagram of an exemplary computer 1700 that maintains the database described above and performs the methods and procedures. Each computer 1700 operates under control of a central
25   processor unit (CPU) 1702, such as a "Pentium" microprocessor and associated integrated circuit chips, available from Intel Corporation of Santa Clara, California, USA. A computer user can input commands and data from a keyboard and display mouse 1704 and can view inputs and computer output at a display 1706. The display is typically a video
30   monitor or flat panel display device. The computer 1700 also includes a direct access storage device (DASD) 1707, such as a fixed hard disk

drive. The memory 1708 typically comprises volatile semiconductor random access memory (RAM). Each computer preferably includes a program product reader 1710 that accepts a program product storage device 1712, from which the program product reader can read data (and

5   to which it can optionally write data). The program product reader can comprise, for example, a disk drive, and the program product storage device can comprise removable storage media such as a magnetic floppy disk, an optical CD-ROM disc, a CD-R disc, a CD-RW disc, or a DVD data disc. If desired, the computers can be connected so they can

10   communicate with each other, and with other connected computers, over a network 1713. Each computer 1700 can communicate with the other connected computers over the network 1713 through a network interface 1714 that enables communication over a connection 1716 between the network and the computer.

15   The computer 1700 operates under control of programming steps that are temporarily stored in the memory 1708 in accordance with conventional computer construction. When the programming steps are executed by the CPU 1702, the pertinent system components perform their respective functions. Thus, the programming steps implement the

20   functionality of the system as described above. The programming steps can be received from the DASD 1707, through the program product reader 1712, or through the network connection 1716. The storage drive 1710 can receive a program product, read programming steps recorded thereon and transfer the programming steps into the memory 1708 for

25   execution by the CPU 1702. As noted above, the program product storage device 1710 can comprise any one of multiple removable media having recorded computer-readable instructions, including magnetic floppy disks and CD-ROM storage discs. Other suitable program product storage devices can include magnetic tape and semiconductor memory

-41-

chips. In this way, the processing steps necessary for operation can be embodied on a program product.

Alternatively, the program steps can be received into the operating memory 1708 over the network 1713. In the network method, the
5   computer receives data including program steps into the memory 1708 through the network interface 1714 after network communication has been established over the network connection 1716 by well-known methods that will be understood by those skilled in the art without further explanation. The program steps are then executed by the CPU 1702 to
10   implement the processing of the Garment Database system.

It should be understood that all of the computers of the system preferably have a construction similar to that shown in Figure 17, so that details described with respect to the Figure 17 computer 1700 will be understood to apply to all computers of the system 1700. This is
15   indicated by multiple computers 1700 shown connected to the network 1713. Any one of the computers 1700 can have an alternative construction, so long as they can communicate with the other computers and support the functionality described herein.

Figure 18 is a flow diagram that illustrates the processing steps
20   performed using the computer illustrated in Figure 17, to maintain and provide access to the databases, such as for identifying polymorphic genetic markers. In particular, the information contained in the database is stored in computers having a construction similar to that illustrated in Figure 17. The first step for maintaining the database, as indicated in
25   Figure 18, is to identify healthy members of a population. As noted above, the population members are subjects that are selected only on the basis of being healthy, and where the subjects are mammals, such as humans, they are preferably selected based upon apparent health and the absence of detectable infections. The step of identifying is represented
30   by the flow diagram box numbered 1802.

-42-

The next step, represented by the flow diagram box numbered 1804, is to obtain identifying and historical information and data relating to the identified members of the population. The information and data comprise parameters for each of the population members, such as

5 member age, ethnicity, sex, medical history, and ultimately genotypic information. Initially, the parameter information is obtained from a questionnaire answered by each member, from whom a body tissue or body fluid sample also is obtained. The step of entering and storing these parameters into the database of the computer is represented by the flow

10 diagram box numbered 1806. As additional information about each population member and corresponding sample is obtained, this information can be inputted into the database and can serve as a sorting parameter.

In the next step, represented by the flow diagram box numbered

15 1808, the parameters of the members are associated with an indexer. This step may be executed as part of the database storage operation, such as when a new data record is stored according to the relational database structure and is automatically linked with other records according to that structure. The step 1806 also may be executed as part

20 of a conventional data sorting or retrieval process, in which the database entries are searched according to an input search or indexing key value to determine attributes of the data. For example, such search and sort techniques may be used to follow the occurrence of known genetic markers and then determine if there is a correlation with diseases for

25 which they have been implicated. Examples of this use are for assessing the frequencies of the p53 and Lipoprotein Lipase polymorphisms.

Such searching of the database also may be valuable for identifying one or more genetic markers whose frequency changes within the population as a function of age, ethnic group, sex, or some other criteria.

30 This can allow the identification of previously unknown polymorphisms

-43-

and, ultimately, identification of a gene or pathway involved in the onset and progression of disease.

In addition, the database can be used for taking an identified polymorphism and ascertaining whether it changes in frequency when the
5   data is sorted according to a selected parameter.

In this way, the databases and methods provided herein permit, among other things, identification of components, particularly key components, of a disease process by understanding its genetic underpinnings, and also an understanding of processes, such as individual
10 drug responses. The databases and methods provided herein also can be used in methods involving elucidation of pathological pathways, in developing new diagnostic assays, identifying new potential drug targets, and in identifying new drug candidates.

Morbidity and/or early mortality associated polymorphisms

15 A database containing information provided by a population of healthy blood donors who were not selected for any particular disease to can be used to identify polymorphisms and the alleles in which they are present, whose frequency decreases with age. These may represent morbidity susceptibility markers and genes.

20 Polymorphisms of the genome can lead to altered gene function, protein function or genome instability. To identify those polymorphisms which have a clinical relevance/utility is the goal of a world-wide scientific effort. It can be expected that the discovery of such polymorphisms will have a fundamental impact on the identification and development of novel
25 drug compounds to cure diseases. However, the strategy to identify valuable polymorphisms is cumbersome and dependent upon the availability of many large patient and control cohorts to show disease association. In particular, genes that cause a general risk of the population to suffer from any disease (morbidity susceptibility genes) will
30 escape these case/control studies entirely.

-44-

Here described is a screening strategy to identify morbidity susceptibility genes underlying a variety of different diseases. The definition of a morbidity susceptibility gene is a gene that is expressed in many different cell types or tissues (housekeeping gene) and its altered

5   function can facilitate the expression of a clinical phenotype caused by disease-specific susceptibility genes that are involved in a pathway specific for this disorder. In other words, these morbidity susceptibility genes predispose people to develop a distinct disease according to their genetic make-up for this disease.

10   Candidates for morbidity susceptibility genes can be found at the bottom level of pathways involving transcription, translation, heat-shock proteins, protein trafficking, DNA repair, assembly systems for subcellular structures (e.g. mitochondria, peroxysomes and other cellular microbodies), receptor signaling cascades, immunology, etc. Those

15   pathways control the quality of life at the cellular level as well as for the entire organism. Mutations/polymorphisms located in genes encoding proteins for those pathways can reduce the fitness of cells and make the organism more susceptible to express the clinical phenotype caused by the action of a disease-specific susceptibility gene. Therefore, these

20   morbidity susceptibility genes can be potentially involved in a whole variety of different complex diseases if not in all. Disease-specific susceptibility genes are involved in pathways that can be considered as disease-specific pathways like glucose-, lipid, hormone metabolism, etc.

The exemplified method permit, among other things, identification

25   of genes and/or gene products involved in a man's general susceptibility to morbidity and/or mortality; use of these genes and/or gene products in studies to elucidate the genetic underpinnings of human diseases; use of these genes and/or gene products in combinatorial statistical analyses without or together with disease-specific susceptibility genes; use of

30   these genes and/or gene products to predict penetrance of disease

-45-

susceptibility genes; use of these genes and/or gene products in predisposition and/or acute medical diagnostics and use of these genes and/or gene products to develop drugs to cure diseases and/or to extend the life span of humans.

5   **SCREENING PROCESS**

The healthy population stratified by age, gender and ethnicity, etc. is a very efficient and a universal screening tool for morbidity associated genes. Changes of allelic frequencies in the young compared to the old population are expected to indicate putative morbidity susceptibility

10   genes. Individual samples of this healthy population base can be pooled to further increase the throughput. In a proof of principle experiment pools of young and old Caucasian females and males were applied to screen more than 400 randomly chosen single nucleotide polymorphisms located in many different genes. Candidate polymorphisms were

15   identified if the allelic difference was greater than 8% between young and old for both or only one of the genders. The initial results were assayed again in at least one independent subsequent experiments. Repeated experiments are necessary to recognize unstable biochemical reactions, which occur with a frequency of about 2-3% and can mimic age-related

20   allelic frequency differences. Average frequency differences and standard deviations are calculated after successful reproducibility of initial results. The final allelic frequency is then compared to a reference population of Caucasian CEPH sample pool. The result should show similar allelic frequencies in the young Caucasian population. Subsequently, the exact

25   allele frequencies of candidates including genotype information were obtained by analyzing all individual samples. This procedure is straight forward with regard to time and cost. It enables the screening of an enormous number of SNPs. So far, several markers with a highly significant association to age were identified and described below.

In general at least 5 individual in a stratified population need to be screened to produce statistically significant results. The frequency of the allele is determined for an age stratified population. Chi square analysis is then performed on the allelic frequencies to determine if the difference

5   between age groups is statistically significant. A p value less than of 0.1 is considered to represent a statistically significant difference. More preferably the p value should be less than 0.05.

Clinical Trials

The identification of markers whose frequency in a population

10   decreases with age also allows for better designed and balanced clinical trials. Currently, if a clinical trial utilizes a marker as a significant endpoint in a study and the marker disappears with age, then the results of the study may be inaccurate. By using methods provided herein, it can be ascertained that if a marker decreases in frequency with age. This

15   information considered and controlled when designing the study. For, example, an age independent marker could be substituted in its place.

The following examples are included for illustrative purposes only and are not intended to limit the scope of the invention.

<div align="center">

EXAMPLE 1

</div>

20   This example describes the use of a database containing information provided by a population of healthy blood donors who were not selected for any particular disease to determine the distribution of allelic frequencies of known genetic markers with age and by sex in a Caucasian subpopulation of the database. The results described in this

25   example demonstrate that a disease-related genetic marker or polymorphism can be identified by sorting a healthy database by a parameter or parameters, such as age, sex and ethnicity.

Generating a database

Blood was obtained by venous puncture from human subjects who met blood bank criteria for donating blood. The blood samples were preserved with EDTA at pH 8.0 and labeled. Each donor provided

5   information such as age, sex, ethnicity, medical history and family medical history. Each sample was labeled with a barcode representing identifying information. A database was generated by entering, for each donor, the subject identifier and information corresponding to that subject into the memory of a computer storage medium using commercially

10 available software, e.g., Microsoft Access.

Model genetic markers

The frequencies of polymorphisms known to be associated at some level with disease were determined in a subpopulation of the subjects represented in the database. These known polymporphisms occur in the

15 p21, p53 and Lipoprotein Lipase genes. Specifically, the N291S polymorphism (N291S) of the Lipoprotein Lipase gene, which results in a substitution of a serine for an asparagine at amino acid codon 291, leads to reduced levels of high density lipoprotein cholesterol (HDL-C) that is associated with an increased risk of males for arteriosclerosis and in

20 particular myocardial infarction (see, Reymer *et al.* (1995) *Nature Genetics 10*:28-34).

The p53 gene encodes a cell cycle control protein that assesses DNA damage and acts as a transcription factor regulating genes that control cell growth, DNA repair and apoptosis (programmed cell death).

25 Mutations in the p53 gene have been found in a wide variety of different cancers, including different types of leukemia, with varying frequency. The loss of normal p53 function results in genomic instability an uncontrolled cell growth. A polymorphism that has been identified in the p53 gene, i.e., the R72P allele, results in the substitution of a proline for

30 an arginine at amino acid codon 72 of the gene.

-48-

The p21 gene encodes a cyclin-dependent kinase inhibitor associated with G1 phase arrest of normal cells. Expression of the p21 gene triggers apoptosis. Polymorphisms of the p21 gene have been associated with Wilms' tumor, a pediatric kidney cancer. One
5   polymorphism of the p21 gene, the S31R polymorphism, results in a substitution of an arginine for a serine at amino acid codon 31.

Database analysis

Sorting of subjects according to specific parameters

The genetic polymorphisms were profiled within segments of the
10   Caucasian subpopulation of the sample bank. For p53 profiling, the genomic DNA isolated from blood from a total of 1277 Caucasian subjects age 18-59 years and 457 Caucasian subjects age 60-79 years was analyzed. For p21 profiling, the genomic DNA isolated from blood from a total of 910 Caucasian subjects age 18-49 years and 824
15   Caucasian subjects age 50-79 years was analyzed. For lipoprotein lipase gene profiling, the genomic DNA from a total of 1464 Caucasian females and 1470 Caucasian males under 60 years of age and a total of 478 Caucasian females and 560 Caucasian males over 60 years of age was analyzed.

20   #### Isolation and analysis of genomic DNA

Genomic DNA was isolated from blood samples obtained from the individuals. Ten milliliters of whole blood from each individual was centrifuged at 2000 x g. One milliliter of the buffy coat was added to 9 ml of 155 mM NH_4Cl, 10 mM $KHCO_3$, and 0.1 mM Na_2EDTA, incubated
25   10 min at room temperature and centrifuged for 10 min at 2000 x g. The supernatant was removed, and the white cell pellet was washed in 155 mM NH_4Cl, 10 mM $KHCO_3$ and 0.1 mM Na_2EDTA and resuspended in 4.5 ml of 50 mM Tris, 5 mM EDTA and 1% SDS. Proteins were precipitated from the cell lysate by 6 mM ammonium acetate, pH 7.3, and then
30   separated from the nucleic acids by centrifugation at 3000 x g. The

-49-

nucleic acid was recovered from the supernatant by the addition of an equal volume of 100% isopropanol and centrifugation at 2000 x g. The dried nucleic acid pellet was hydrated in 10 mM Tris, pH 7.6, and 1 mM Na$_2$EDTA and stored at 4° C.

5   Assays of the genomic DNA to determine the presence or absence of the known genetic markers were developed using the BiomassPROBE™ detection method (primer oligo base extension) reaction. This method uses a single detection primer followed by an oligonucleotide extension step to give products, which can be readily resolved by mass

10   spectrometry, and, in particular, MALDI-TOF mass spectrometry. The products differ in length depending on the presence or absence of a polymorphism. In this method, a detection primer anneals adjacent to the site of a variable nucleotide or sequence of nucleotides and the primer is extended using a DNA polymerase in the presence of one or more

15   dideoxyNTPs and, optionally, one or more deoxyNTPs. The resulting products are resolved by MALDI-TOF mass spectrometry. The mass of the products as measured by MALDI-TOF mass spectrometry makes possible the determination of the nucleotide(s) present at the variable site.

       First, each of the Caucasian genomic DNA samples was subjected

20   to nucleic acid amplification using primers corresponding to sites 5' and 3' of the polymorphic sites of the p21 (S31R allele), p53 (R72P allele) and Lipoprotein Lipase (N291S allele) genes. One primer in each primer pair was biotinylated to permit immobilization of the amplification product to a solid support. Specifically, the polymerase chain reaction primers used

25   for amplification of the relevant segments of the p21, p53 and lipoprotein lipase genes are shown below: US4p21c31-2F (SEQ ID NO: 9) and US5p21-2R (SEQ ID NO: 10) for p21 gene amplification; US4-p53-ex4-F (also shown as p53-ex4US4 (SEQ ID NO: 2)) and US5-p53/2-4R (also shown as US5P53/4R (SEQ ID NO: 3)) for p53 gene amplification; and

US4-LPL-F2 (SEQ ID NO: 16) and US5-LPL-R2 (SEQ ID NO: 17) for lipoprotein lipase gene amplification.

Amplification of the respective DNA sequences was conducted according to standard protocols. For example, primers may be used in a concentration of 8 pmol. The reaction mixture (e.g., total volume 50 μl) may contain Taq-polymerase including 10x buffer and dTNPs. Cycling conditions for polymerase chain reaction amplification may typically be initially 5 min. at 95°C, followed by 1 min. at 94°C, 45 sec at 53°C, and 30 sec at 72°C for 40 cycles with a final extension time of 5 min at 72°C. Amplification products may be purified by using Qiagen's PCR purification kit (No. 28106) according to manufacturer's instructions. The elution of the purified products from the column can be done in 50 μl TE-buffer (10mM Tris, 1 mM EDTA, pH 7.5).

The purified amplification products were immobilized via a biotin-avidin linkage to streptavidin-coated beads and the double-stranded DNA was denatured. A detection primer was then annealed to the immobilized DNA using conditions such as, for example, the following: 50 μl annealing buffer (20 mM Tris, 10 mM KCl, 10 mM $(NH_4)_2SO_4$, 2 mM $MgSO_2$, 1% Triton X-100, pH 8) at 50°C for 10 min, followed by washing of the beads three times with 200 μl washing buffer (40 mM Tris, 1 mM EDTA, 50 mM NaCl, 0.1% Tween 20, pH 8.8) and once in 200 μl TE buffer.

The PROBE extension reaction was performed, for example, by using some components of the DNA sequencing kit from USB (No. 70770) and dNTPs or ddNTPs from Pharmacia. An exemplary protocol could include a total reaction volume of 45 μl, containing of 21 μl water, 6 μl Sequenase-buffer, 3 μl 10 mM DTT solution, 4.5 μl, 0.5 mM of three dNTPs, 4.5 μl, 2 mM the missing one ddNTP, 5.5 μl glycerol enzyme dilution buffer, 0.25 μl Sequenase 2.0, and 0.25 pyrophosphatase. The reaction can then by pipetted on ice and incubated for 15 min at room temperature and for 5 min at 37°C. The beads may be washed three

times with 200 μl washing buffer and once with 60 μl of a 70 mM NH$_4$-Citrate solution.

The DNA was denatured to release the extended primers from the immobilized template. Each of the resulting extension products was

5   separately analyzed by MALDI-TOF mass spectrometry using 3-hydroxypicolinic acid (3-HPA) as matrix and a UV laser.

Specifically, the primers used in the PROBE reactions are as shown below: P21/31-3 (SEQ ID NO: 12) for PROBE analysis of the p21 polymorphic site; P53/72 (SEQ ID NO: 4) for PROBE analysis of the p53

10 polymorphic site; and LPL-2 for PROBE analysis of the lipoprotein lipase gene polymorphic site. In the PROBE analysis of the p21 polymorphic site, the extension reaction was performed using dideoxy-C. The products resulting from the reaction conducted on a "wild-type" allele template (wherein codon 31 encodes a serine) and from the reaction

15 conducted on a polymorphic S31R allele template (wherein codon 31 encodes an arginine) are shown below and designated as P21/31-3 Ser (wt) (SEQ ID NO: 13) and P21/31-3 Arg (SEQ ID NO: 14), respectively. The masses for each product as can be measured by MALDI-TOF mass spectrometry are also provided (i.e., 4900.2 Da for the wild-type product

20 and 5213.4 Da for the polymorphic product).

In the PROBE analysis of the p53 polymorphic site, the extension reaction was performed using dideoxy-C. The products resulting from the reaction conducted on a "wild-type" allele template (wherein codon 72 encodes an arginine) and from the reaction conducted on a polymorphic

25 R72P allele template (wherein codon 72 encodes a proline) are shown below and designated as Cod72 G Arg (wt) and Cod72 C Pro, respectively. The masses for each product as can be measured by MALDI-TOF mass spectrometry are also provided (i.e., 5734.8 Da for the wild-type product and 5405.6 Da for the polymorphic product).

In the PROBE analysis of the lipoprotein lipase gene polymorphic site, the extension reaction was performed using a mixture of ddA and ddT. The products resulting from the reaction conducted on a "wild-type" allele template (wherein codon 291 encodes an asparagine) and

5 from the reaction conducted on a polymorphic N291S allele template (wherein codon 291 encodes a serine) are shown below and designated as 291Asn and 291Ser, respectively. The masses for each product as can be measured by MALDI-TOF mass spectrometry are also provided (i.e., 6438.2 Da for the wild-type product and 6758.4 Da for the

10 polymorphic product).

P53-1 (R72P)

PCR Product length: 407 bp (SEQ ID NO: 1)

```
                                       US4-p53-ex4-F
                                   ctg aggacctggt cctctgactg
15  ctcttttcac ccatctacag tccccttgc cgtcccaagc aatggatgat ttgatgctgt
    ccccggacga tattgaacaa tggttcactg aagacccagg tccagatgaa gctcccagaa
     P53/72              72R
    tgccagaggc tgctccccgc gtggccctg caccagcagc tcctacaccg gcggcccctg
                     c 72P
20  caccagcccc ctcctggccc ctgtcatctt ctgtcccttc ccagaaaacc taccagggca
    gctacggttt ccgtctgggc ttcttgcatt ctgggacagc caagtctgtg acttgcacgg
    tcagttgccc tgaggggctg gcttccatga gacttcaa
                             US5-p53/2-4R
```

Primers (SEQ ID NOs: 2-4)

25 p53-ex4FUS4   ccc agt cac gac gtt gta aaa cgc tga gga cct ggt cct ctg ac

US5P53/4R   agc gga taa caa ttt cac aca ggt tga agt ctc atg gaa gcc

P53/72   gcc aga ggc tgc tcc cc

Masses

Allele	Product Termination: ddC	SEQ #	Length	Mass
P53/72	gccagaggctgctcccc	5	17	5132.4
Cod72 G Arg (wt)	gccagaggctgctccccgc	6	19	5734.8
Cod72 C Pro	gccagaggctgctcccc	7	18	5405.6

Biotinylated US5 primer is used in the PCR amplification.

35 **LPL-1 (N291S)**

Amino acid exchange asparagine to serine at codon 291 of the lipoprotein lipase gene.

-53-

PCR Product length: 251 bp (SEQ ID NO: 15)

US4-LPL-F2 (SEQ ID NO: 16)

```
gcgctccatt catctcttca tcgactctct gttgaatgaa gaaaatccaa gtaaggccta
caggtgcagt tccaaggaag cctttgagaa agggctctgc ttgagttgta gaaagaaccg
          LPL-2                291N
ctgcaacaat ctgggctatg agatcaataa agtcagagcc aaaagaagca gcaaaatgta
                               g 291S
cctgaagact cgttctcaga tgccc
               US4-LPL-R2
```

Primers (SEQ ID NOs: 16-18):

US4-LPL-F2 ccc agt cac gac gtt gta aaa cgg cgc tcc att cat ctc ttc

US5-LPL-R2 agc gga taa caa ttt cac aca ggg ggc atc tga gaa cga gtc

LPL-2   caa tct ggg cta tga gat ca

Masses

Allele	Product Termination: ddA, ddT	SEQ #	Length	Mass
LPL-2	caatctgggctatgagatca	19	20	6141
291 Asn	caatctgggctatgagatcaa	20	21	6438.2
291 Ser	caatctgggctatgagatcagt	21	22	6758.4

Biotinylated US5 primer is used in the PCR amplification.

P21-1 (S31R)

Amino acid exchange serine to arginine at codon 31 of the tumor suppressor gene p21. Product length: 207 bp (SEQ ID NO: 8)

```
US4p21c31-2F
                              gtcc gtcagaaccc atgcggcagc
                                   p21/31-3 31S
aaggcctgcc gccgcctctt cggcccagtg gacagcgagc agctgagccg cgactgtgat
                                                      a 31R
gcgctaatgg cgggctgcat ccaggaggcc cgtgagcgat ggaacttcga ctttgtcacc
gagacaccac tggaggg
          US5p21-2R
```

Primers (SEQ ID NOs: 9-11)

US4p21c31-2F ccc agt cac gac gtt gta aaa cgg tcc gtc aga acc cat gcg g

US5p21-2R   agc gga taa caa ttt cac aca ggc tcc agt ggt gtc tcg tgt ac

P21/31-3   cag cga gca gct gag

-54-

Masses

Allele	Product Termination: ddC	SEQ #	Length	Mass
p21/31-3	cagcgagcagctgag	12	15	4627
P21/31-3 Ser (wt)	cagcgagcagctgagc	13	16	4900.2
P21/31-3 Arg	cagcgagcagctgagac	14	17	5213.4

Biotinylated US5 primer is used in the PCR amplification.

Each of the Caucasian subject DNA samples was individually analyzed by MALDI-TOF mass spectrometry to determine the identity of the nucleotide at the polymorphic sites. The genotypic results of each assay can be entered into the database. The results were then sorted according to age and/or sex to determine the distribution of allelic frequencies by age and/or sex. As depicted in the Figures showing histograms of the results, in each case, there was a differential distribution of the allelic frequencies of the genetic markers for the p21, p53 and lipoprotein lipase gene polymorphisms.

Figure 8 shows the results of the p21 genetic marker assays reveals a statistically significant decrease (from 13.3% to 9.2%) in the frequency of the heterozygous genotype (S31R) in Caucasians with age (18-49 years of age compared to 50-79 years of age). The frequencies of the homozygous (S31 and R31) genotypes for the two age groups are also shown, as are the overall frequencies of the S31 and R31 alleles in the two age groups (designated as ˙S31 and ˙R31, respectively in the Figure).

Figures 7A-C shows the results of the p53 genetic marker assays and reveals a statistically significant decrease (from 6.7% to 3.7%) in the frequency of the homozygous polymorphic genotype (P72) in Caucasians with age (18-59 years of age compared to 60-79 years of age). The frequencies of the homozygous "wild-type" genotype (R72) and the heterozygous genotype (R72P) for the two age groups are also shown, as are the overall frequencies of the R72 and P72 alleles in the two age

-55-

groups (designated as *R72 and *P72, respectively in the Figure). These results are consistent with the observation that allele is not benign, as p53 regulates expression of a second protein, p21, which inhibits cyclin-dependent kinases (CDKs) needed to drive cells through the

5 cell-cycle (a mutation in either gene can disrupt the cell cycle leading to increased cell division).

Figure 2C shows the results of the lipoprotein lipase gene genetic marker assays reveals a statistically significant decrease (from 1.97% to 0.54%) in the frequency of the polymorphic allele (S291) in Caucasian

10 males with age (see also Reymer *et al.* (1995) *Nature Genetics* 10:28-34). The frequencies of this allele in Caucasian females of different age groups are also shown.

EXAMPLE 2

This example describes the use of MALDI-TOF mass spectrometry

15 to analyze DNA samples of a number of subjects as individual samples and as pooled samples of multiple subjects to assess the presence or absence of a polymorphic allele (the 353Q allele) of the Factor VII gene and determine the frequency of the allele in the group of subjects. The results of this study show that essentially the same allelic frequency can

20 be obtained by analyzing pooled DNA samples as by analyzing each sample separately and thereby demonstrate the quantitative nature of MALDI-TOF mass spectrometry in the analysis of nucleic acids.

Factor VII

Factor VII is a serine protease involved in the extrinsic blood

25 coagulation cascade. This factor is activated by thrombin and works with tissue factor (Factor III) in the processing of Factor X to Factor Xa. There is evidence that supports an association between polymorphisms in the Factor VII gene and increased Factor VII activity which can result in an elevated risk of ischemic cardiovascular disease, including myocardial

30 infarction. The polymorphism investigated in this study is R353Q (i.e., a

substitution of a glutamic acid residue for an arginine residue at codon 353 of the Factor VII gene) (see Table 5).

Analysis of DNA samples for the presence or absence of the 353Q allele of the Factor VII gene

5   Genomic DNA was isolated from separate blood samples obtained from a large number of subjects divided into multiple groups of 92 subjects per group. Each sample of genomic DNA was analyzed using the BiomassPROBE™ assay as described in Example 1 to determine the

10   presence or absence of the 353Q polymorphism of the Factor VII gene.

First, DNA from each sample was amplified in a polymerase chain reaction using primers F7-353FUS4 (SEQ ID NO: 24) and F7-353RUS5 (SEQ ID NO: 26) as shown below and using standard conditions, for example, as described in Example 1. One of the primers was biotinylated

15   to permit immobilization of the amplification product to a solid support. The purified amplification products were immobilized via a biotin-avidin linkage to streptavidin-coated beads and the double-stranded DNA was denatured. A detection primer was then annealed to the immobilized DNA using conditions such as, for example, described in Example 1. The

20   detection primer is shown as F7-353-P (SEQ ID NO: 27) below. The PROBE extension reaction was carried out using conditions, for example, such as those described in Example 1. The reaction was performed using ddG.

The DNA was denatured to release the extended primers from the

25   immobilized template. Each of the resulting extension products was separately analyzed by MALDI-TOF mass spectrometry. A matrix such as 3-hydroxypicolinic acid (3-HPA) and a UV laser could be used in the MALDI-TOF mass spectrometric analysis. The products resulting from the reaction conducted on a "wild-type" allele template (wherein codon 353

30   encodes an arginine) and from the reaction conducted on a polymorphic 353Q allele template (wherein codon 353 encodes a glutamic acid) are

-57-

shown below and designated as 353 CGG and 353 CAG, respectively. The masses for each product as can be measured by MALDI-TOF mass spectrometry are also provided (i.e., 5646.8 Da for the wild-type product and 5960 Da for the polymorphic product).

5   The MALDI-TOF mass spectrometric analyses of the PROBE reactions of each DNA sample were first conducted separately on each sample (250 nanograms total concentration of DNA per analysis). The allelic frequency of the 353Q polymorphism in the group of 92 subjects was calculated based on the number of individual subjects in which it was

10 detected.

Next, the samples from 92 subjects were pooled (250 nanograms total concentration of DNA in which the concentration of any individual DNA is 2.7 nanograms) and the pool of DNA was subjected to MALDI-TOF mass spectrometric analysis. The area under the signal

15 corresponding to the mass of the 353Q polymorphism PROBE extension product in the resulting spectrum was integrated in order to quantitate the amount of DNA present. The ratio of this amount to total DNA was used to determine the allelic frequency of the 353Q polymorphism in the group of subjects. This type of individual sample vs. pooled sample analysis

20 was repeated for numerous different groups of 92 different samples.

The frequencies calculated based on individual MALDI-TOF mass spectrometric analysis of the 92 separate samples of each group of 92 are compared to those calculated based on MALDI-TOF mass spectrometric analysis of pools of DNA from 92 samples in Figure 9.

25 These comparisons are shown as "pairs" of bar graphs in the Figure, each pair being labeled as a separate "pool" number, e.g., P1, P16, P2, etc. Thus, for example, for P1, the allelic frequency of the polymorphism calculated by separate analysis of each of the 92 samples was 11.41% and the frequency calculated by analysis of a pool of all of the 92 DNA

30 samples was 12.09%.

The similarity in frequencies calculated by analyzing separate DNA samples individually and by pooling the DNA samples demonstrates that it is possible, through the quantitative nature of MALDI-TOF mass spectrometry, to analyze pooled samples and obtain accurate frequency

5 determinations. The ability to analyze pooled DNA samples significantly reduces the time and costs involved in the use of the non-selected, healthy databases as described herein. It has also been shown that it is possible to decrease the DNA concentration of the individual samples in a pooled mixture from 2.7 nanograms to 0.27 nanograms without any

10 change in the quality of the spectrum or the ability to quantitate the amount of sample detected.

Factor VII R353Q PROBE Assay

PROBE Assay for cod353 CGG>CAG (Arg>Gln), Exon 9 G>A.

PCR fragment: 134 bp (incl. US tags; SEQ ID Nos. 22 and 23)

15 Frequency of A allele: Europeans about 0.1, Japanese/Chinese about 0.03-0.05 (Thromb. Haemost. 1995, 73:617-22; Diabetologia 1998, 41:760-6):

```
                          F7-353FUS4>
      1201 GTGCCGGCTA CTCGGATGGC AGCAAGGACT CCTGCAAGGG GGACAGTGGA
20    GGCCCACACATG
           F7-353-P>        A                <F7-353RUS5
      1261 CCACCCACTA CCGGGGCACG TGGTACCTGA CGGGCATCGT CAGCTGGGGC
      CAGGGCTGCG
```

```
Primers  (SEQ ID NOs.24-26)                                          Tm⁰³
25    F7-353FUS4   CCC AGT CAC GAC GTT GTA AAA CGA TGG CAG CAA GGA CTC CTG   64°C
      F7-353-P     CAC ATG CCA CCC ACT ACC
      F7-353RUS5   AGC GGA TAA CAA TTT CAC ACA GGT GAC GAT GCC CGT CAG GTA C   64°C
```

Masses

Allele	Product Termination: ddG	SEQ #	Length	Mass
F7-353-P	atgccacccactacc	27	18	5333.6
353 CGG	cacatgccacccactaccg	28	19	5646.8
353 CAG	cacatgccacccactaccag	29	20	5960
US5-bio bio-	agcggataacaatttcacacagg	30	23	7648.6

Conclusion

The above examples demonstrate an effect of altered frequency of disease causing genetic factors within the general population. Interpretation of those results allows prediction of the medical relevance

5 of polymorphic genetic alterations. In addition, conclusions can be drawn with regard to their penetrance, diagnostic specificity, positive predictive value, onset of disease, most appropriate onset of preventive strategies, and the general applicability of genetic alterations identified in isolated populations to panmixed populations. Therefore, an age- and sex-stratified

10 population-based sample bank that is ethnically homogenous is a suitable tool for rapid identification and validation of genetic factors regarding their potential medical utility.

<div align="center">EXAMPLE 3</div>

MORBIDITY AND MORTALITY MARKERS

15 **Sample Band and Initial Screening**

Healthy samples were obtained through the blood bank of San Bernardino, CA. Donors signed prior to the blood collection a consent form and agreed that their blood will be used in genetic studies with regard to human aging. All samples were anomymized. Tracking back of

20 samples is not possible.

Isolation of DNA from blood samples of a healthy donor population

Blood is obtained from a donor by venous puncture and preserved with 1mM EDTA pH 8.0. Ten milliliters of whole blood from each donor was centrifuged at 2000x g. One milliliter of the buffy coat was added to

25 9 milliters of 155mM NH_4Cl, 10mM $KHCO_3$, and 0.1mM Na_2EDTA, incubated 10 minutes at room temperature and centrifuged for 10 minutes at 2000x g. The supernatant was removed, and the white cell pellet was washed in 155mM NH_4Cl, 10mM $KHCO_3$, and 0.1mM Na_2EDTA and resuspended in 4.5 milliliters of 50mM Tris, 5mM EDTA,

30 and 1% SDS. Proteins were precipitated from the cell lysate by 6M

Ammonium Acetate, pH 7.3, and separated from the nucleic acid by centrifugation 3000x g. The nucleic acid was recovered from the supernatant by the addition of an equal volume of 100% isopropanol and centrifugation at 2000x g. The dried nucleic acid pellet was hydrated in 5 IOmM Tris pH 7.6 and 1mM Na2EDTA and stored at 4C.

In this study, samples were pooled as shown in Table 1. Both parents of the blood donors were of Caucasian origin.

Table 1

Pool ID	Sex	Age-range	# individuals
SP1	Female	18-39 years	276
SP2	Males	18-39 years	276
SP3	Females	60-69 years	184
SP4	Males	60-79 years	368

More than 400 SNPs were tested using all four pools. After one test run 34 assays were selected to be re-assayed at least once. Finally, 10 assays showed repeatedly differences in allele frequencies of several percent and, therefore, fulfilled the criteria to be tested using the individual samples. Average allele frequency and standard deviation is tabulated in Table 2.

Table 2

Assay ID	SP1	SP1-STD	SP2	SP2-STD	SP3	SP3-STD	SP4	SP4-STD
47861	0.457	0.028	0.433	0.042	0.384	0.034	0.006	0.015
47751	0.276	0.007	0.483	0.006	0.428	0.052	0.400	0.097
48319	0.676	0.013	0.627	0.018	0.755	0.006	0.006	0.034
48070	0.581	0.034	0.617	0.045	0.581	n.a.	0.539	0.032
49807	0.504	0.034	0.422	0.020	0.477	0.030	0.556	0.005

49534	0.537	0.017	0.503	n.a.	0.623	0.023	0.535	0.009
49733	0.560	0.006	0.527	0.059	0.546	0.032	0.436	0.016
49947	0.754	0.008	0.763	0.047	0.736	0.052	0.689	0.025
50128	0.401	0.022	0.363	0.001	0.294	0.059	0.345	0.013
63306	0.697	0.012	0.674	0.013	0.712	0.017	0.719	0.005

So far, 7 out of the 10 potential morbidity markers were fully analyzed. Additional information about genes in which these SNPs are located was gathered through publicly databases like Genbank.

AKAPS

Candidate morbidity and mortality markers include housekeeping genes, such as genes involved in signal transduction. Among such genes are the A-kinase anchoring proteins (AKAPs) genes, which participate in signal transduction pathways involving protein phosphorylation. Protein phosphorylation is an important mechanism for enzyme regulation and the transduction of extracellular signals across the cell membrane in eukaryotic cells. A wide variety of cellular substrates, including enzymes, membrane receptors, ion channels and transcription factors, can be phosphorylated in response to extracellular signals that interact with cells. A key enzyme in the phosphorylation of cellular proteins in response to hormones and neurotransmitters is cyclic AMP (cAMP)-dependent protein kinase (PKA). Upon activation by cAMP, PKA thus mediates a variety of cellular responses to such extracellular signals. An array of PKA isozymes are expressed in mammalian cells. The PKAs usually exist as inactive tetramers containing a regulatory (R) subunit dimer and two catalytic (C) subunits. Genes encoding three C subunits (Cα, Cβ and Cγ) and four R subunits (RIα, RIβ, RIIα and RIIβ) have been identified [see Takio *et al.* (1982) *Proc. Natl. Acad. Sci. U.S. A. 79*:2544-2548; Lee *et al.* (1983)

-62-

Proc. Natl. Acad. Sci. U.S.A. 80:3608-3612; Jahnsen *et al.* (1996) *J. Biol. Chem. 261*:12352-12361; Clegg *et al.* (1988) *Proc. Natl. Acad. Sci. U.S.A. 85*:3703-3707; and Scott (1991) *Pharmacol. Ther. 50*:123-145]. The type I (RI) *α* and type II (RII) *α* subunits are distributed ubiquitously,

5   whereas RI*β* and RII*β* are present mainly in brain [see. *e.g.*, Miki and Eddy (1999) *J. Biol. Chem. 274*:29057-29062]. The type I PKA holoenzyme (RI*α* and RI*β*) is predominantly cytoplasmic, whereas the majority of type II PKA (RII*α* and RII*β*) associates with cellular structures and organelles [Scott (1991) *Pharmacol. Ther. 50*:123-145]. Many hormones and other

10   signals act through receptors to generate cAMP which binds to the R subunits of PKA and releases and activates the C subunits to phosphorylate proteins. Because protein kinases and their substrates are widely distributed throughout cells, there are mechanisms in place in cells to localize protein kinase-mediated responses to different signals. One

15   such mechanism involves subcellular targeting of PKAs through association with anchoring proteins, referred to as A-kinase anchoring proteins (AKAPs), that place PKAs in close proximity to specific organelles or cytoskeletal components and particular substrates thereby providing for more specific PKA interactions and localized responses [see,

20   *e.g.*, Scott *et al.* (1990) *J. Biol. Chem. 265*:21561-21566; Bregman *et al.* (1991) *J. Biol. Chem. 266*:7207-7213; and Miki and Eddy (1999) *J. Biol. Chem. 274*:29057-29062]. Anchoring not only places the kinase close to preferred substrates, but also positions the PKA holoenzyme at sites where it can optimally respond to fluctuations in the second messenger

25   cAMP [Mochly-Rosen (1995) *Science 268*:247-251; Faux and Scott (1996) *Trends Biochem. Sci. 21*:312-315; Hubbard and Cohen (1993) *Trends Biochem. Sci. 18*:172-177].

Up to 75% of type II PKA is localized to various intracellular sites through association of the regulatory subunit (RII) with AKAPs [see, *e.g.*,

30   Hausken *et al.* (1996) *J. Biol. Chem. 271*:29016-29022]. RII subunits of

PKA bind to AKAPs with nanomolar affinity [Carr *et al.* (1992) *J. Biol. Chem. 267*:13376-13382], and many AKAP-RII complexes have been isolated from cell extracts. RI subunits of PKA bind to AKAPs with only micromolar affinity [Burton *et al.* (1997) *Proc. Natl. Acad. Sci. U.S.A.*

5   *94*:11067-11072]. Evidence of binding of a PKA RI subunit to an AKAP has been reported [Miki and Eddy (1998) *J. Biol. Chem 273*:34384-34390] in which RIα-specific and RIα/RIIα dual specificity PKA anchoring domains were identified on FSC1/AKAP82. Additional dual specific AKAPs, referred to as D-AKAP1 and D-AKAP2, which interact with the

10 type I and type II regulatory subunits of PKA have also been reported [Huang *et al.* (1997) *J. Biol. Chem. 272*:8057-8064; Huang *et al.* (1997) *Proc. Natl. Acad. Sci. U.S.A. 94*:11184-11189].

More than 20 AKAPs have been reported in different tissues and species. Complementary DNAs (cDNAs) encoding AKAPs have been

15 isolated from diverse species, ranging from *Caenorhabditis elegans* and *Drosophilia* to human [see, *e.g.*, Colledge and Scott (1999) *Trends Cell Biol. 9*:216-221]. Regions within AKAPs that mediate association with RII subunits of PKA have been identified. These regions of approximately 10-18 amino acid residues vary substantially in primary sequence, but

20 secondary structure predictions indicate that they are likely to form an amphipathic helix with hydrophobic residues aligned along one face of the helix and charged residues along the other [Carr *et al.* (1991) *J. Biol. Chem. 266*:14188-14192; Carr *et al.* (1992) *J. Biol. Chem. 267*:13376-13382]. Hydrophobic amino acids with a long aliphatic side chain, e.g.,

25 valine, leucine or isoleucine, may participate in binding to RII subunits [Glantz *et al.* (1993) *J. Biol. Chem. 268*:12796-12804].

Many AKAPs also have the ability to bind to multiple proteins, including other signaling enzymes. For example, AKAP79 binds to PKA, protein kinase C (PKC) and the protein phosphatase calcineurin (PP2B)

30 [Coghlan *et al.* (1995) *Science 267*:108-112 and Klauck *et al.* (1996)

-64-

Science 271:1589-1592]. Therefore, the targeting of AKAP79 to neuronal postsynaptic membranes brings together enzymes with opposite catalytic activities in a single complex.

AKAPs thus serve as potential regulatory mechanisms that increase
5   the selectivity and intensity of a cAMP-mediated response. There is a need, therefore, to identify and elucidate the structural and functional properties of AKAPs in order to gain a complete understanding of the important role these proteins play in the basic functioning of cells.

AKAP10

10   The sequence of a human AKAP10 cDNA (also referred to as D-AKAP2) is available in the GenBank database, at accession numbers AF037439 (SEQ ID NO: 31) and NM 007202. The AKAP10 gene is located on chromosome 17.

The sequence of a mouse D-AKAP2 cDNA is also available in the
15   GenBank database (see accession number AF021833). The mouse D-AKAP2 protein contains an RGS domain near the amino terminus that is characteristic of proteins that interact with Gα subunits and possess GTPase activating protein-like activity [Huang *et al.* (1997) *Proc. Natl. Acad. Sci. U.S.A. 94*:11184-11189]. The human AKAP10 protein also
20   has sequences homologous to RGS domains. The carboxy-terminal 40 residues of the mouse D-AKAP2 protein are responsible for the interaction with the regulatory subunits of PKA. This sequence is fairly well conserved between the mouse D-AKAP2 and human AKAP10 proteins.

Polymorphisms of the human AKAP10 gene and polymorphic AKAP10
25   **proteins**

Polymorphisms of AKAP genes that alter gene expression, regulation, protein structure and/or protein function are more likely to have a significant effect on the regulation of enzyme (particularly PKA) activity, cellular transduction of signals and responses thereto and on the
30   basic functioning of cells than polymorphisms that do not alter gene

and/or protein function. Included in the polymorphic AKAPs provided herein are human AKAP10 proteins containing differing amino acid residues at position number 646.

Amino acid 646 of the human AKAP10 protein is located in the
5 carboxy-terminal region of the protein within a segment that participates in the binding of R-subunits of PKAs. This segment includes the carboxy-terminal 40 amino acids.

The amino acid residue reported for position 646 of the human AKAP10 protein is an isoleucine. Polymorphic human AKAP10 proteins
10 provided herein have the amino acid sequence but contain residues other than isoleucine at amino acid position 646 of the protein. In particular embodiments of the polymorphic human AKAP10 proteins provided herein, the amino acid at position 646 is a valine, leucine or phenylalanine residue.

15 **An A to G transition at nucleotide 2073 of the human AKAP10 coding sequence**

As described herein, an allele of the human AKAP10 gene that contains a specific polymorphism at position 2073 of the coding sequence and thereby encodes a valine at position 646 has been detected
20 in varying frequencies in DNA samples from younger and older segments of the human population. In this allele, the A at position 2073 of the AKAP10 gene coding sequence is changed from an A to a G, giving rise to an altered sequence in which the codon for amino acid 646 changes from ATT, coding for isoleucine, to GTT, coding for valine.

25 **Morbidity marker 1: human protein kinase A anchoring protein (AKAP10-1)**

PCR Amplification and BiomassPROBE assay detection of AKAP10-1 in a healthy donor population

PCR Amplification of donor population for AKAP 10

30 PCR primers were synthesized by OPERON using phosphoramidite chemistry. Amplification of the AKAP10 target sequence was carried out

in single 50μl PCR reaction with 100ng-1ug of pooled human genomic DNAs in a 50μl PCR reaction. Individual DNA concentrations within the pooled samples were present in equal concentration with the final concentration ranging from 1-25ng. Each reaction containing IX PCR

5   buffer (Qiagen, Valencia, CA), 2OOuM dNTPs, 1U Hotstar Taq polymerase (Qiagen, Valencia, CA), 4mM MgCl$_2$, and 25pmol of the forward primer containing the universal primer sequence and the target specific sequence 5'-TCTCAATCATGTGCATTGAGG-3'(SEQ ID NO: 45), 2pmol of the reverse primer

10   5'-AGCGGATAACAATTTCACACAGGGATCACACAGCCATCAGCAG-3' (SEQ ID NO: 46), and IOpmol of a biotinylated universal primer complementary to the 5' end of the PCR amplicon 5'-AGCGGATAACAATTTCACACAGG-3'(SEQ ID NO: 47). After an initial round of amplification with the target with the specific forward and

15   reverse primer, the 5' biotinylated universal primer then hybridized and acted as a reverse primer thereby introducing a 3' biotin capture moiety into the molecule. The amplification protocol results in a 5'-biotinylated double stranded DNA amplicon and dramatically reduces the cost of high throughput genotyping by eliminating the need to 5' biotin label each

20   forward primer used in a genotyping. Thermal cycling was performed in 0.2mL tubes or 96 well plate using an MJ Research Thermal Cycler (calculated temperature) with the following cycling parameters: 94° C for 5 min; 45 cycles: 94° C for 20 sec, 56° C for 30 sec, 72° C for 60 sec; 72° C 3min.

25   <u>Immobilization of DNA</u>

The 50μl PCR reaction was added to 25ul of streptavidin coated magnetic bead (Dynal) prewashed three times and resuspended in 1M NH$_4$Cl, 0.06M NH$_4$OH. The PCR amplicons were allowed to bind to the beads for 15 minutes at room temperature. The beads were then collected with a

30   magnet and the supernatant containing unbound DNA was removed. The

-67-

unbound strand was release from the double stranded amplicons by incubation in 100mM NaOH and washing of the beads three times with 10mM Tris pH 8.0.

BiomassPROBE assay analysis of donor population for AKAP10-1 (clone 48319)

Genotyping using the BiomassPROBE assay methods was carried out by resuspending the DNA coated magnetic beads in 26mM Tris-HCl pH 9.5, 6.5 mM MgCl$_2$ and 50mM each of dTTP and 50mM each of ddCTP, ddATP, ddGTP, 2.5U of a thermostable DNA polymerase (Ambersham) and 20pmol of a template specific oligonucleotide PROBE primer 5'-CTGGCGCCCACGTGGTCAA-3' (SEQ ID NO: 48) (Operon). Primer extension occurs with three cycles of oligonucleotide primer hybridization and extension. The extension products were analyzed after denaturation from the template with 50mM NH$_4$Cl and transfer of 150nL each sample to a silicon chip preloaded with 150nL of H3PA matrix material. The sample material was allowed to crystallize and was analyzed by MALDI-TOF (Bruker, PerSeptive). The SNP that is present in AKAP10-1 is a T to C transversion at nucleotide number 156277 of the sequence of a genomic clone of the AKAP10 gene (GenBank Accession No. AC005730) (SEQ ID NO: 36). SEQ ID NO: 35: represents the nucleotide sequence of human chromosome 17, which contains the genomic nucleotide sequence of the human AKAP10 gene, and SEQ ID NO: represents the nucleotide sequence of human chromosome 17, which contains the genomic nucleotide sequence of the human AKAP10-1 allele. The mass of the primer used in the BioMass probe reaction was 5500.6 daltons. In the presence of the SNP, the primer is extended by the addition of ddC, which has a mass of 5773.8. The wildtype gene results in the addition of dT and ddG to the primer to produce an extension product having a mass of 6101 daltons.

-68-

The frequency of the SNP was measured in a population of age selected healthy individuals. Five hundred fifty-two (552) individuals between the ages of 18-39 years (276 females, 276 males) and 552 individuals between the ages of 60-79 (184 females between the ages of
5   60-69, 368 males between the age of 60-79) were tested for the presence of the polymorphism localized in the non-translated 3'region of AKAP 10. Differences in the frequency of this polymorphism with increasing age groups were observed among healthy individuals. Statistical analysis showed that the significance level for differences in
10 the allelic frequency for alleles between the "younger" and the "older" populations was p = 0.0009 and for genotypes was p = 0.003. Differences between age groups are significant. For the total population allele significance is p = 0.0009, and genotype significance is p = 0.003.

This marker led to the best significant result with regard to allele
15 and genotype frequencies in the age-stratified population. Figure 19 shows the allele and genotype frequency in both genders as well as in the entire population. For latter the significance for alleles was p = 0.0009 and for genotypes was p = 0.003. The young and old populations were in Hardy-Weinberg equilibrium. A preferential change of one particular
20 genotype was not seen.

The polymorphism is localized in the non-translated 3'-region of the gene encoding the human protein kinase A anchoring protein (AKAP10). The gene is located on chromosome 17. Its structure includes 15 exons and 14 intervening sequences (introns). The encoded protein is
25 responsible for the sub-cellular localization of the cAMP-dependent protein kinase and, therefore, plays a key role in the G-protein mediated receptor-signaling pathway (Huang et al. PNAS (1007) 94:11184-11189). Since its localization is outside the coding region, this polymorphism is most likely in linkage disequilibrium (LD) with other non-synonymous
30 polymorphisms that could cause amino acid substitutions and

subsequently alter the function of the protein. Sequence comparison of different Genbank database entries concerning this gene revealed further six potential polymorphisms of which two are supposed to change the respective amino acid (see Table 3).

5   Table 3

Exon	Codon	Nucleotides	Amino acid
3	100	GCT > GCC	Ala > Ala
4	177	AGT > GTG	Met > Val
8	424	GGG > GGC	Gly > Gly
10	524	CCG > CTG	Pro > Leu
12	591	GTG > GTC	Val > Val
12	599	CGC > CGA	Arg > Arg

Morbitity marker 2: human protein kinase A anchoring protein
15   **(AKAP10-5)**

Discovery of AKAP10-5 Allele (SEQ ID NO: 33)

Genomic DNA was isolated from blood (as described above) of seventeen (17) individuals with a genotype CC at the AKAP10-1 gene locus and a single heterozygous individual (CT) (as described). A target
20   sequence in the AKAP10-1 gene which encodes the C-terminal PKA binding domain was amplified using the polymerase chain reaction. PCR primers were synthesized by OPERON using phosphoramidite chemistry. Amplification of the AKAP10-1 target sequence was carried out in individual 50μl PCR reaction with 25ng of human genomic DNA
25   templates. Each reaction containing I X PCR buffer (Qiagen, Valencia, CA), 200μM dNTPs, IU Hotstar Taq polymerase (Qiagen, Valencia, CA), 4mM $MgCl_2$, 25pmol of the forward primer (Ex13F) containing the universal primer sequence and the target specific sequence 5'-TCC CAA AGT GCT GGA ATT AC-3' (SEQ ID NO: 53), and 2pmol of the reverse
30   primer (Ex14R) 5'-GTC CAA TAT ATG CAA ACA GTT G-3' (SEQ ID NO:

-70-

54). Thermal cycling was performed in 0.2mL tubes or 96 well plate using an MJ Research Thermal Cycler (MJ Research, Waltham, MA) (calculated temperature) with the following cycling parameters: 94° C for 5 min; 45 cycles; 94° C for 20 sec, 56° C for 30 sec, 72° C for 60 sec;

5   72° C 3min. After amplification the amplicons were purified using a chromatography (Mo Bio Laboratories (Solana Beach, CA)).

The sequence of the 18 amplicons, representing the target region, was determined using a standard Sanger cycle sequencing method with 25nmol of the PCR amplicon, 3.2uM DNA sequencing primer 5'-CCC ACA

10 GCA GTT AAT CCT TC-3'(SEQ ID NO: 55), and chain terminating dRhodamine labeled 2', 3' dideoxynucleotides (PE Biosystems, Foster City, CA) using the following cycling parameters: 96° C for 15 seconds; 25 cycles: 55° C for 15 seconds, 60° C for 4 minutes. The sequencing products precipitated by 0.3M NaOAc and ethanol. The precipitate was

15 centrifuged and dried. The pellets were resuspended in deionized formamide and separated on a 5% polyacrylimide gel. The sequence was determined using the "Sequencher" software (Gene Codes, Ann Arbor, MI).

The sequence of all 17 of the amplicons, which are homozygous

20 for the AKAP10-1 SNP of the amplicons, revealed a polymorphism at nucleotide position 152171 (numbering for GenBank Accession No. AC005730 for AKAP10 genomic clone (SEQ ID NO: 35)) with A replaced by G. This SNP can also be designated as located at nucleotide 2073 of a cDNA clone of the wildtype AKAP10 (GenBank Accession No.

25 AF037439) (SEQ ID NO: 31). The amino acid sequence of the human AKAP10 protein is provided as SEQ ID NO: 32. This single nucleotide polymorphism was designated as AKAP10-5 (SEQ ID NO: 33) and resulted in a substitution of a valine for an isoleucine residue at amino acid position 646 of the amino acid sequence of human AKAP10 (SEQ ID

30 NO: 32).

-71-

PCR Amplification and BiomassPROBE assay detection of AKAP10-5 in a healthy donor population

The healthy population stratified by age is a very efficient and a universal screening tool for morbidity associated genes by allowing for the detection of changes of allelic frequencies in the young compared to the
5   old population. Individual samples of this healthy population base can be pooled to further increase the throughput.

Healthy samples were obtained through the blood bank of San Bernardino, CA. Both parents of the blood donors were of Caucasian
10   origin. Practically a healthy subject, when human, is defined as human donor who passes blood bank criteria to donate blood for eventual use in the general population. These criteria are as follows: free of detectable viral, bacterial, mycoplasma, and parasitic infections; not anemic; and then further selected based upon a questionnaire regarding history (see
15   Figure 3). Thus, a healthy population represents an unbiased population of sufficient health to donate blood according to blood bank criteria, and not further selected for any disease state. Typically such individuals are not taking any medications.

PCR primers were synthesized by OPERON using phosphoramidite
20   chemistry. Amplification of the AKAP10 target sequence was carried out in a single $50\mu l$ PCR reaction with 100ng- $1\mu g$ of pooled human genomic DNAs in a $50\mu l$ PCR reaction. Individual DNA concentrations within the pooled samples were present in equal concentration with the final concentration ranging from 1-25ng. Each reaction contained 1X PCR
25   buffer (Qiagen, Valencia, CA), $200\mu M$ dNTPs, 1U Hotstar Taq polymerase (Qiagen, Valencia, CA), 4mM $MgCl_2$, and 25pmol of the forward primer containing the universal primer sequence and the target specific sequence 5'-AGCGGATAACAATTTCACACAGGGAGCTAGCTTGGAAGAT TGC-3' (SEQ ID NO: 41), 2pmol of the reverse primer
30   5'-GTCCAATATATGCAAACAGTTG-3' (SEQ ID NO: 54), and 10pmol of a

-72-

biotinylated universal primer complementary to the 5' end of the PCR amplicon BIO:5'-AGCGGATAACAATTTCACACAGG-3' (SEQ ID NO: 43). After an initial round of amplification with the target with the specific forward and reverse primer, the 5' biotinylated universal primer can then

5 be hybridized and acted as a forward primer thereby introducing a 5' biotin capture moiety into the molecule. The amplification protocol resulted in a 5'-biotinylated double stranded DNA amplicon and dramatically reduced the cost of high throughput genotyping by eliminating the need to 5' biotin label every forward primer used in a

10 genotyping.

Themal cycling was performed in 0.2mL tubes or 96 well plate using an MJ Research Thermal Cycler (calculated temperature) with the following cycling parameters: 94° C for 5 min; 45 cycles: 94° C for 20 sec, 56° C for 30 sec; 72° C for 60 sec; 72° C 3min.

15 **Immobilization of DNA**

The 50 μl PCR reaction was added to 25μL of streptavidin coated magnetic beads (Dynal, Oslo, Norway), which were prewashed three times and resuspended in 1M NH_4Cl, 0.06M NH_4OH. The 5' end of one strand of the double stranded PCR amplicons were allowed to bind to the

20 beads for 15 minutes at room temperature. The beads were then collected with a magnet and the supernatant containing unbound DNA was removed. The hybridized but unbound strand was released from the double stranded amplicons by incubation in 100mM NaOH and washing of the beads three times with 10mM Tris pH 8.0.

25 **Detection of AKAP10-5 using BiomassPROBE™ Assay**

BiomassPROBE™ assay of primer extension analysis (see, U.S. Patent No. 6,043,031) of donor population for AKAP 10-5 (SEQ ID NO: 33) was performed. Genotyping using these methods was carried out by resuspending the DNA coated magnetic beads in 26mM Tris-HCL pH 9.5,

30 6.5 mM $MgCl_2$, 50mM dTTP, 50mM each of ddCTP, ddATP, ddGTP, 2.5U

-73-

of a thermostable DNA polymerase (Ambersham), and 20pmol of a
template specific oligonucleotide PROBE primer
5'-ACTGAGCCTGCTGCATAA-3' (SEQ ID NO: 44) (Operon). Primer
extension occurs with three cycles of oligonucleotide primer with
5   hybridization and extension. The extension products were analyzed after
denaturation from the template with 50 mM NH$_4$Cl and transfer of 150 nL
of each sample to a silicon chip preloaded with 150 nl of H3PA matrix
material. The sample material was allowed to crystallize and analyzed by
MALDI-TOF (Bruker, PerSeptive). The primer has a mass of 5483.6
10   daltons. The SNP results in the additional of a ddC to the primer, giving a
mass of 5756.8 daltons for the extended product. The wild type results in
the addition a T and ddG to the primer giving a mass of 6101 daltons.

The frequency of the SNP was measured in a population of age
selected healthy individuals. Seven hundred thirteen (713) individuals
15   under 40 years of age (360 females, 353 males) and 703 individuals over
60 years of age (322 females, 381 males) were tested for the presence of
the SNP, AKAP10-5 (SEQ ID NO: 33). Results are presented below in
Table 1.

TABLE 1 AKAP10-5 (2073V) frequency comparison in 2 age groups					
			<40	>60	delta G allele
Female	Alleles	*G	38.6	34.6	4.0
		*A	61.4	65.4	
	Genotypes	G	13.9	11.8	2.1
		GA	49.4	45.7	
		A	36.7	42.5	
Male	Alleles	*G	41.4	37.0	4.4
		*A	58.6	63.0	
	Genotypes	G	18.4	10.8	7.7

-74-

		GA	45.9	52.5	
		A	35.7	36.7	
Total	Alleles	*G	40.0	35.9	4.1
		*A	60.0	64.1	
	Genotypes	G	16.1	11.2	4.9
		GA	47.7	49.4	
		A	36.2	39.4	

Figure 20 graphically shows these results of allele and genotype distribution in the age and sex stratified Caucasian population.

Morbidity marker 3: human methionine sulfoxide reductase A (msrA)

The age-related allele and genotype frequency of this marker in both genders and the entire population is shown in Figure 21. The decrease of the homozygous CC genotype in the older male population is highly significant.

Methionine sulfoxide reductase A (#63306)

PCR Amplification and BiomassPROBE assay detection of the human methioine sulfoxid reductase A (h-msr-A) in a healthy donor population

PCR Amplification of donor population for h-msr-A

PCR primers were synthesized by OPERON using phosphoramidite chemistry. Amplification of the AKAP10 target sequence was carried out in single 50μl PCR reaction with 100ng-1ug of pooled human genomic DNA templates in a 50μl PCR reaction. Individual DNA concentrations within the pooled samples were present in an equal concentration with the final concentration ranging from 1-25ng. Each reaction containing I X PCR buffer (Qiagen, Valencia, CA), 200μM dNTPs, 1U Hotstar Taq polymerase (Qiagen, Valencia, CA), 4mM MgCl$_2$, 25pmol of the forward primer containing the universal primer sequence and the target specific sequence 5'-TTTCTCTGCACAGAGAGGC-3' (SEQ ID NO: 49), 2pmol of

-75-

the reverse primer

5'-AGCGGATAACAATTTCACACAGGGCTGAAATCCTTCGCTTTACC-3'

(SEQ ID NO: 50), and 10pmol of a biotinylated universal primer

complementary to the 5' end of the PCR amplicon

5   5'-AGCGGATAACAATTTCACACAGG-3' (SEQ ID NO: 51). After an initial

round of amplification of the target with the specific forward and reverse

primers, the 5' biotinylated universal primer was then hybridized and

acted as a reverse primer thereby introducing a 3' biotin capture moiety

into the molecule. The amplification protocol results in a 5'-biotinylated

10 double stranded DNA amplicon and and dramatically reduces the cost of

high throughput genotyping by eliminating the need to 5' biotin label each

forward primer used in a genotyping. Thermal cycling was performed in

0.2mL tubes or 96 well plate using an MJ Research Thermal Cycler

(calculated temperature) with the following cycling parameters: 94° C for

15 5 min; 45 cycles: 94° C for 20 sec, 56° C for 30 sec, 72° C for 60 sec;

72° C 3min.

Immobilization of DNA

     The 50μl PCR reaction was added to 25ul of streptavidin coated

magnetic bead (Dynal) prewashed three times and resuspended in 1M

20 NH$_4$Cl, 0.06M NH$_4$OH. The PCR amplicons were allowed to bind to the

beads for 15 minutes at room temperature. The beads were then

collected with a magnet and the supernatant containing unbound DNA

was removed. The unbound strand was release from the double stranded

amplicons by incubation in 100mM NaOH and washing of the beads three

25 times with 10mM Tris pH 8.0.

BiomassPROBE assay analysis of donor population for h-msr A

     Genotyping using the BiomassPROBE assay methods was carried

out by resuspending the he DNA coated magnetic beads in 26mM

Tris-HCl pH 9.5, 6.5 mM MgCl$_2$, 50mM of dTTPs and 50mM each of

30 ddCTP, ddATP, ddGTP, 2.5U of a thermostable DNA polymerase

-76-

(Ambersham), and 20pmol of a template specific oligonucleotide PROBE primer 5'-CTGAAAAGGGAGAGAAAG-3' (Operon) (SEQ ID NO: 52). Primer extension occurs with three cycles of oligonucleotide primer with hybridization and extension. The extension products were analyzed after denaturation from the template with 50mM NH_4Cl and transfer of 150nl each sample to a silicon chip preloaded with 150nl of H3PA matrix material. The sample material was allowed to crystallize and analyzed by MALDI-TOF (Bruker, PerSeptive). The SNP is represented as a T to C tranversion in the sequence of two ESTs. The wild type is represented by having a T at position 128 of GenBank Accession No. AW 195104, which represents the nucleotide sequence of an EST which is a portion of the wild type human msrA gene (SEQ ID NO: 39). The SNP is presented as a C at position 129 of GenBank Accession No. AW 874187, which represents the nucleotide sequence of an EST which is a portion of an allele of the human msrA gene (SEQ ID NO: 40).

In a genomic sequence the SNP is represented as an A to G transversion. The primer utilized in the BioMass probe reaction had a mass of 5654.8 daltons. In the presence of the SNP the primer is extended by the incorporation of a ddC and has a mass of 5928. In the presence of the wildtype the primer is extended by adding a dT and a DDC to produce a mass of 6232.1 daltons.

The frequency of the SNP was measured in a population of age selected healthy individuals. Five hundred fifty-two (552) individuals between the ages of 18-39 years (276 females, 276 males and 552 individuals between the age of 60-79 (184 females between the ages of 60-69, 368 males between the age of 60-79) were tested for the presence of the polymorphism localized in the nontranslated 3'region of h-msr-A.

Genotype difference between male age group among healthy individuals is significant. For the male population allele significance is

-77-

p = 0.0009 and genotype significance is p = 0.003. The age-related allele and genotype frequency of this marker in both genders and the entire population is shown in Figure 21. The decrease of the homozygous CC genotype in the older male population is highly significant.

5   The polymorphism is localized in the non-translated 3'-region of the gene encoding the human methionine sulfoxide reductase (h-msrA). The exact localization is 451 base pairs downstream the stop codon (TAA). It is very likely that this SNP is in linkage disequilibrium (LD) with another polymorphism more upstream in the coding or promoter region; thus, it is

10   not directly cause morbidity. The enzyme methionine sulfoxide reductase has been proposed to exhibit multiple biological functions. It may serve to repair oxidative protein damage but also play an important role in the regulation of proteins by activation or inactivation of their biological functions (Moskovitz et al. (1990) PNAS 95:14071-14075). It has also

15   been shown that its activity is significantly reduced in brain tissues of Alzheimer patients (Gabbita et al., (1999) J. Neurochem 73:1660-1666). It is scientifically conceivable that proteins involved in the metabolism of reactive oxygen species are associated to disease.

CONCLUSION

20   The use of the healthy population provides for the identification of morbidity markers. The identification of proteins involved in the G-protein coupled signaling transduction pathway or in the detoxification of oxidative stress can be considered as convincing results. Further confirmation and validation of other potential polymorphisms already

25   identified in *silico* in the gene encoding the human protein kinase A anchoring protein could even provide stronger association to morbidity and demonstrate that this gene product is a suitable pharmaceutical or diagnostic target.

EXAMPLE 4

MALDI-TOF Mass Spectrometry Analysis

All of the products of the enzyme assays listed below were analyzed by MALDI-TOF mass spectrometry. A diluted matrix solution
5 (0.15μL) containing of 10:1 3-hydroxypicolinic acid:ammonium citrate in 1:1 water:acetonitrile diluted 2.5-fold with water was pipetted onto a SpectroChip (Sequenom, Inc.) and was allowed to crystallize. Then, 0.15μL of sample was added. A linear PerSeptive Voyager DE mass spectrometer or Bruker Biflex MALDI-TOF mass spectrometer, operating in
10 positive ion mode, was used for the measurements. The sample plates were kept at 18.2 kV for 400 nm after each UV laser shot (approximate 250 laser shots total), and then the target voltage was raised to 20 kV. The original spectra were digitized at 500 MHz.

EXAMPLE 5

15 **Sample Conditioning**

Where indicated in the examples below, the products of the enzymatic digestions were purified with ZipTips (Millipore, Bedford, MA). The ZipTips were pre-wetted with 10 μL 50% acetonitrile and equilibrated 4 times with 10 μl 0.1 M TEAAc. The oligonucleotide fragments were
20 bound to the C18 in the ZipTip material by continuous aspiration and dispension of each sample into the ZipTip. Each digested oligonucleotide was conditioned by washing with 10 μL 0.1 M TEAAc, followed by 4 washing steps with 10 μL H$_2$O. DNA fragments were eluted from the Ziptip with 7 μL 50% acetonitrile.

25 Any method for condition the samples may be employed. Methods for conditioning, which generally is used to increase peak resolution, are well known (see, *e.g.*, International PCT application No. WO 98/20019).

EXAMPLE 6

DNA Glycosylase-Mediated Sequence Analysis

DNA Glycosylases modifies DNA at each position that a specific nucleobase resides in the DNA, thereby producing abasic sites. In a
5   subsequent reaction with another enzyme, a chemical, or heat, the phosphate backbone at each abasic site can be cleaved.

The glycosylase utilized in the following procedures was uracil-DNA glycosylase (UDG). Uracil bases were incorporated into DNA fragments in each position that a thymine base would normally occupy by amplifying a
10   DNA target sequence in the presence of uracil. Each uracil substituted DNA amplicon was incubated with UDG, which cleaved each uracil base in the amplicon, and was then subjected to conditions that effected backbone cleavage at each abasic site, which produced DNA fragments. DNA fragments were subjected to MALDI-TOF mass spectrometry
15   analysis. Genetic variability in the target DNA was then assessed by analyzing mass spectra.

Glycosylases specific for nucleotide analogs or modified nucleotides, as described herein, can be substituted for UDG in the following procedures. The glycosylase methods described hereafter, in
20   conjunction with phosphate backbone cleavage and MALDI, can be used to analyze DNA fragments for the purposes of SNP scanning, bacteria typing, methylation analysis, microsatellite analysis, genotyping, and nucleotide sequencing and re-sequencing.

A.   Genotyping

25   A glycosylase procedure was used to genotype the DNA sequence encoding UCP-2 (Uncoupling Protein 2). The sequence for UCP-2 is deposited in GenBank under accession number AF096289. The sequence variation genotyped in the following procedure was a cytosine (C-allele) to thymine (T-allele) variation at nucleotide position 4790, which results in a
30   alanine to valine mutation at position 55 in the UCP-2 polypeptide.

-80-

DNA was amplified using a PCR procedure with a 50 μL reaction volume containing of 5 pmol biotinylated primer having the sequence 5'-TGCTTATCCCTGTAGCTACCCTGTCTTGGCCTTGCAGATCCAA-3' (SEQ ID NO: 91), 15 pmol non-biotinylated primer having the sequence 5'-

5 AGCGGATAACAATTTCACACAGGCCATCACACCGCGGTACTG-3' (SEQ ID NO: 92), 200 μM dATP, 200 μM dCTP, 200 μM dGTP, 600 μM dUTP (to fully replace dTTP), 1.5 mM to 3 mM MgCl$_2$, 1 U of HotStarTaq polymerase, and 25 ng of CEPH DNA. Amplification was effected with 45 cycles at an annealing temperature of 56°C.

10 The amplification product was then immobilized onto a solid support by incubating 50 μL of the amplification reaction with 5 μL of prewashed Dynabeads for 20 minutes at room temperature. The supernatant was removed, and the beads were incubated with 50 μL of 0.1 M NaOH for 5 minutes at room temperature to denature the double-

15 stranded PCR product in such a fashion that single-stranded DNA was linked to the beads. The beads were then neutralized by three washes with 50 μL 10 mM TrisHCl (pH 8). The beads were resuspended in 10 μL of a 60mM TrisHCl/1mM EDTA (pH 7.9) solution, and 1 U uracil DNA glycosylase was added to the solution for 45 minutes at 37°C to remove

20 uracil nucleotides present in the single-stranded DNA linked to the beads. The beads were then washed two times with 25 μL of 10 mM TrisHCl (pH 8) and once with 10 μL of water. The biotinylated strands were then eluted from the beads with 12 μL of 2 M NH$_4$OH at 60°C for 10 minutes. The backbone of the DNA was cleaved by incubating the samples for 10

25 min at 95°C (with a closed lid), and ammonia was evaporated from the samples by incubating the samples for 11 min at 80°C.

The cleavage fragments were then analyzed by MALDI-TOF mass spectrometry as described in Example 4. The T-allele generated a unique fragment of 3254 Daltons. The C-allele generated a unique fragment of

30 4788 Daltons. These fragments were distinguishable in mass spectra.

-81-

Thus, the above-identified procedure was successfully utilized to genotype individuals heterozygous for the C-allele and T-allele in UCP-2.

B.   Glycosylase Analysis Utilizing Pooled DNA Samples

The glycosylase assay was conducted using pooled samples to

5   detect genetic variability at the UCP-2 locus. DNA of known genotype was pooled from eleven individuals and was diluted to a fixed concentration of 5 ng/μL. The procedure provided in Example 3A was followed using 2 pmol of forward primer having a sequence of 5'-CCCAGTCACGACGTTGTAAAACGTCTTGGCCTTGCAGATCCAAG- 3'

10   (SEQ ID NO: 93) and 15 pmol of reverse primer having the sequence 5'-AGCGGATAACAATTTCACACAGGCCATCACACCGCGGTACTG-3' (SEQ ID NO: 94). In addition, 5 pmol of biotinylated primer having the sequence 5'bioCCCAGTCACGACGTTGTAAAACG 3' (SEQ ID NO: 97) may be introduced to the PCR reaction after about two cycles. The

15   fragments were analyzed via MALDI-TOF mass spectroscopy (Example 4). As determined in Example 3A, the T-allele, which generated a unique fragment of 3254 Daltons, could be distinguished in mass spectra from the C-allele, which generated a unique fragment of 4788 Daltons.   Allelic frequency in the pooled samples was quantified by integrating the area

20   under each signal corresponding to an allelic fragment. Integration was accomplished by hand calculations using equations well known to those skilled in the art. In the pool of eleven samples, this procedure suggested that 40.9% of the individuals harbored the T allele and 59.09% of the individuals harbored the C allele.

25   **C.   Glycosylase-Mediated Microsatellite Analysis**

A glycosylase procedure was utilized to identify microsatellites of the Bradykinin Receptor 2 (BKR-2) sequence. The sequence for BKR-2 is deposited in GenBank under accession number X86173. BKR-2 includes a SNP in the promoter region, which is a C to T variation, as well as a

30   SNP in a repeated unit, which is a G to T variation. The procedure

provided in Example 3A was utilized to identify the SNP in the promotor region, the SNP in the microsattelite repeat region, and the number of repeated units in the microsattelite region of BKR-2. Specifically, a forward PCR primer having the sequence 5'-

5   CTCCAGCTGGGCAGGAGTGC-3' (SEQ ID NO: 95) and a reverse primer having the sequence 5'-CACTTCAGTCGCTCCCT-3' (SEQ ID NO: 96) were utilized to amplify BKR-2 DNA in the presence of uracil. The amplicon was fragmented by UDG followed by backbone cleavage. The cleavage fragments were analyzed by MALDI-TOF mass spectrometry as

10   described in Example 4.

With regard to the SNP in the BKR-2 promotor region having a C to T variation, the C-allele generated a unique fragment having a mass of 7342.4 Daltons and the T-allele generated a unique fragment having a mass of 7053.2 Daltons. These fragments were distinguishable in mass

15   spectra. Thus, the above-identified procedure was successfully utilized to genotype individuals heterozygous for the C-allele and T-allele in the promotor region of BKR-2.

With regard to the SNP in the BKR-2 repeat region having a G to T variation, the T-allele generated a unique fragment having a mass of 1784

20   Daltons, which was readily detected in a mass spectrum. Hence, the presence of the T-allele was indicative of the G to T sequence variation in the repeat region of BKR-2.

In addition, the number of repeat regions was distinguished between individuals having two repeat sequences and individuals having

25   three repeat sequences in BKR-2. The DNA of these individuals did not harbor the G to T sequence variation in the repeat sequence as each repeat sequence contained a G at the SNP locus. The number of repeat regions was determined in individual samples by calculating the area under a signal corresponding to a unique DNA fragment having a mass of

30   2771.6 Daltons. This signal in spectra generated from individuals having

two repeat regions had an area that was thirty-three percent less than the area under the same signal in spectra generated from individuals having three repeat regions. Thus, the procedures discussed above can be utilized to genotype individuals for the number of repeat sequences

5 present in BKR-2.

D.   Bisulfite Treatment Coupled with Glycosylase Digestion

Bisulfite treatment of genomic DNA can be utilized to analyze positions of methylated cytosine residues within the DNA. Treating nucleic acids with bisulfite deaminates cytosine residues to uracil

10 residues, while methylated cytosine remains unmodified. Thus, by comparing the sequence of a PCR product generated from genomic DNA that is not treated with bisulfite with the sequence of a PCR product generated from genomic DNA that is treated with bisulfite, the degree of methylation in a nucleic acid as well as the positions where cytosine is

15 methylated can be deduced.

Genomic DNA (2 μg) was digested by incubation with 1 μL of a restriction enzyme at 37°C for 2 hours. An aliquot of 3 M NaOH was added to yield a final concentration of 0.3M NaOH in the digestion solution. The reaction was incubated at 37°C for 15 minutes followed by

20 treatment with 5.35M urea, 4.44M bisulfite, and 10mM hydroquinone, where the final concentration of hydroquinone is 0.5 mM.

The sample that was treated with bisulfite (sample A) was compared to the same digestion sample that had not undergone bisulfite treatment (sample B). After sample A was treated with bisulfite as

25 described above, sample A and sample B were amplified by a standard PCR procedure. The PCR procedure included the step of overlaying each sample with mineral oil and then subjecting the sample to thermocycling (20 cycles of 15 minutes at 55°C followed by 30 seconds at 95°C). The PCR reaction contained four nucleotide bases, C, A, G, and U. The

30 mineral oil was removed from each sample, and the PCR products were

-84-

purified with glassmilk. Sodium iodide (3 volumes) and glassmilk (5 μL) were added to samples A and B. The samples were then placed on ice for 8 minutes, washed with 420 μL cold buffer, centrifuged for 10 seconds, and the supernatant fractions were removed. This process was

5 repeated twice and then 25 μL of water was added. Samples were incubated for 5 minutes at 37 °C, were centrifuged for 20 seconds, and the supernatant fraction was collected, and then this incubation/centrifugation/supernatant fraction collection procedure was repeated. 50 μL 0.1 M NaOH was then added to the samples to denature

10 the DNA. The samples were incubated at room temperature for 5 minutes, washed three times with 50 μL of 10 mM TrisHCl (pH 8), and resuspended in 10 μL 60mM TrisHCl/1mM EDTA, pH 7.9.

The sequence of PCR products from sample A and sample B were then treated with 2U of UDG (MBI Fermentas) and then subjected to

15 backbone cleavage, as described herein. The resulting fragments from each of sample A and sample B were analyzed by MALDI-TOF mass spectroscopy as described in Example 4. Sample A gave rise to a greater number of fragments than the number of fragments arising from sample B, indicative that the nucleic acid harbored at least one methylated

20 cytosine moiety.

EXAMPLE 7

Fen-Ligase-Mediated Haplotyping

Haplotyping procedures permit the selection of a fragment from one of an individual's two homologous chromosomes and to genotype linked SNPs

25 on that fragment. The direct resolution of haplotypes can yield increased information content, improving the diagnosis of any linked disease genes or identifying linkages associated with those diseases. In previous studies, haplotypes were typically reconstructed indirectly through pedigree analysis (in cases where pedigrees were available) through

laborious and unreliable allele-specific PCR or through single-molecule dilution methods well known in the art.

A haplotyping procedure was used to determine the presence of two SNPs, referred to as SNP1 and SNP2, located on one strand in a DNA

5   sample. The haplotyping procedure used in this assay utilized Fen-1, a site-specific "flap" endonuclease that cleaves DNA "flaps" created by the overlap of two oligonucleotides hybridized to a target DNA strand. The two overlapping oligonucleotides in this example were short arm and long arm allele-specific adaptors. The target DNA was an amplified nucleic

10   acid that had been denatured and contained SNP1 and SNP2.

The short arm adaptor included a unique sequence not found in the target DNA. The 3' distal nucleotide of the short arm adaptor was identical to one of the SNP1 alleles. Moreover, the long arm adaptor included two regions: a 3' region complementary to the short arm and a

15   5'gene-specific region complementary to the fragment of interest adjacent to the SNP. If there was a match between the adaptor and one of the homologues, the Fen enzyme recognized and cleaved the overlapping flap. The short arm of the adaptor was then ligated to the remainder of the target fragment (minus the SNP site). This ligated fragment was used

20   as the forward primer for a second PCR reaction in which only the ligated homologue was amplified. The second PCR product (PCR2) was then analyzed by mass spectrometry. If there was no match between the adaptors and the target DNA, there was no overlap, no cleavage by Fen-1, and thus no PCR2 product of interest.

25   If there was more than one SNP in the sequence of interest, the second SNP (SNP2) was found by using an adaptor that was specific for SNP2 and hybridizing the adaptor to the PCR2 product containing the first SNP. The Fen-ligase and amplification procedures were repeated for the PCR2 product containing the first SNP. If the amplified product yielded a

30   second SNP, then SNP1 and SNP2 were on the same fragment.

If the SNP is unknown, then four allele-specific adaptors (e.g. C, G, A, and T) can be used to hybridize with the target DNA. The substrates are then treated with the Fen-ligase protocol, including amplification. The PCR2 products may be analyzed by PROBE, as described herein, to

5 determine which adaptors were hybridized to the DNA target and thus identify the SNPs in the sequence.

A Fen-ligase assay was used to detect two SNPs present in Factor VII. These SNPs are located 814 base pairs apart from each other. SNP1 was located at position 8401 (C to T), and SNP2 was located at 9215 (G

10 to A) (SEQ ID #).

A.   First Amplification Step

A PCR product (PCR1) was generated for a known heterozygous individual at SNP1, a short distance from the 5' end of the SNP. Specifically, a 10 μL PCR reaction was performed by mixing 1.5 mM

15 MgCl$_2$, 200 μM of each dNTP, 0.5 U HotStar polymerase, 0.1μM of a forward primer having the sequence 5'-GCG CTC CTG TCG GTG CCA (SEQ ID NO: 56), 0.1μM of a reverse primer having the sequence 5'-GCC TGA CTG GTG GGG CCC (SEQ ID NO: 57), and 1 ng of genomic DNA. The annealing temperature was 58°C, and the amplification process

20 yielded fragments that were 861 bp in length.

The PCR1 reaction mixture was divided in half and was treated with an exonuclease 1/SAP mixture (0.22 μL mixture/5 μL PCR1 reaction) which contained 1.0μL SAP and 0.1 μL exon1. The exonuclease treatment was done for 30 minutes at 37°C and then 20 minutes at

25 85°C to denature the DNA.

B.   Adaptor Oligonucleotides

A solution of allele-specific adaptors (C and T), containing of one long and one short oligonucleotide per adaptor, was prepared. The long arm and short arm oligonucleotides of each adaptor (10μM) were mixed in

30 a 1:1 ratio and heated for 30 seconds at 95°C. The temperature was

reduced in 2°C increments to 37°C for annealing. The C-adaptor had a short arm sequence of 5'-CAT GCA TGC ACG GTC (SEQ ID NO: 58) and a long arm sequence of 5'-CAG AGA GTA CCC CTC GAC CGT GCA TGC ATG (SEQ ID NO: 59). Hence, the long arm of the adaptor was 30 bp

5   (15 bp gene-specific), and the short arm was 15bp. The T-adaptor had a short arm sequence of 5'-CAT GCA TGC ACG GTT (SEQ ID NO: 60) and a long arm sequence of 5'-GTA CGT ACG TGC CAA CTC CCC ATG AGA GAC (SEQ ID NO: 61). The adaptor could also have a hairpin structure in which the short and long arm are separated by a loop containing of 3 to

10 10 nucleotides (SEQ ID NO: 118).

C. FEN-ligase reaction

In two tubes (one tube for each allele-specific adaptor per sample) was placed a solution (Solution A) containing of 3.5 μl 10 mM 16%PEG/50 mM MOPS, 1.2 μl 25 mM MgCl$_2$, 1.5 μl 10X Ampligase

15 Buffer, and 2.5 μl PCR1. Each tube containing Solution A was incubated at 95°C for 5 minutes to denature the PCR1 product. A second solution (Solution B) containing of 1.65 μl Ampligase (Thermostable ligase, Epicentre Technologies), 1.65 μl 200ng/μl MFEN (from *Methanococcus jannaschii*), and 3.0 μl of an allele specific adaptor (C or T) was prepared.

20 Thus, different variations of Solution B, each variation containing of different allele-specific adaptors, were made. Solution B was added to Solution A at 95°C and incubated at 55°C for 3 hours. The total reaction volume was 15.0 μl per adaptor-specific reaction. For a bi-allelic system, 2 x 15.0 μl reactions were required.

25 The Fen-ligase reaction in each tube was then deactivated by adding 8.0 μl 10 mM EDTA. Then, 1.0 μl exoIII/Buffer (70%/30%) solution was added to each sample and incubated 30 minutes at 37°C, 20 minutes at 70°C (to deactivate exoIII), and 5 minutes at 95°C (to denature the sample and dissociate unused adaptor from template). The

30 samples were cooled in an ice slurry and purified on UltraClean PCR

Clean-up (MoBio) spin columns which removed all fragments less than 100 base pairs in length. The fragments were eluted with 50 μl H$_2$0.

D. Second Amplification Step

     A second amplification reaction (PCR2) was conducted in each

5 sample tube using the short arm adaptor (C or T) sequence as the forward primer (minus the SNP1 site). Only the ligated homologue was amplified. A standard PCR reaction was conducted with a total volume of 10.0 μl containing of 1X Buffer (final concentration), 1.5 mM final concentration MgCl$_2$, 200 μM final concentration dNTPs, 0.5 U HotStar polymerase, 0.1

10 μM final concentration forward primer 5'-CAT GCA TGC ACG GT (SEQ ID NO: 62), 0.1μM final concentration reverse primer 5'-GCC TGA CTG GTG GGG CCC (SEQ ID NO: 63), and 1.0 μl of the purified FEN-ligase reaction solution. The annealing temperature was 58°C. The PCR2 product was analyzed by MALDI TOF mass spectroscopy as described in Example 4.

15 The mass spectrum of Fen SNP1 showed a mass of 6084.08 Daltons, representing the C allele.

E.   Genotyping Additional SNPs

     The second SNP (SNP2) can be found by using an adaptor that is specific for SNP2 and hybridizing that adaptor to the PCR2 product

20 containing the first SNP. The Fen-ligase and amplification procedures are repeated for the PCR2 product containing the first SNP. If the amplified product yields a second SNP, then SN1 and SN2 are on the same fragment. The mass spectrum of SNP2, representing the T allele, showed a mass of 6359.88 Daltons.

25   This assay can also be performed upon pooled DNA to yield haplotype frequencies as described herein. The Fen-ligase assay can be used to analyze multiplexes as described herein.

EXAMPLE 8

Nickase-Mediated Sequence Analysis

A DNA nickase, or DNase, was used to recognize and cleave one strand
of a DNA duplex. Two nickases usd were NY2A nickase and NYS1
5   nickase (Megabase) which cleave DNA at the following sites:

NY2A: 5'...R AG...3'

3'...Y↓TC...5' where R = A or G and Y = C or T

NYS1: 5'...↓CC[A/G/T]...3'

3'... GG[T/C/A]...5'.

10   **A.   Nickase Digestion**

Tris-HCl (10 mM), KCl (10 mM, pH 8.3), magnesium acetate (25
mM), BSA (1 mg/mL), and 6 U of Cvi NY2A or Cvi NYS1 Nickase
(Megabase Research) were added to 25 pmol of double-stranded
oligonucleotide template having a sequence of 5'-CGC AGG GTT CCC
15   TCG TCG CAC TGG GCA TGT G-3' (SEQ ID NO: 90, Operon, Alameda,
CA) synthesized using standard phosphoramidite chemistry . With a total
volume of 20μL, the reaction mixture was incubated at 37°C for 5 hours,
and the digestion products were purified using ZipTips (Millipore, Bedford,
MA) as described in Example 5. The samples were analyzed by MALTY-
20   TOM mass spectroscopy as described in Example 1. The nickase Cvi
NY2A yielded three fragments with masses 4049.76 Daltons, 5473.14
Daltons, and 9540.71 Daltons. The Cvi NYS1 nickase yielded fragments
with masses 2063.18 Daltons, 3056.48 Daltons, 6492.81 Daltons, and
7450.14 Daltons.

25   **B.   Nickase Digestion of Pooled Samples**

DQA (HLA ClassII-DQ Alpha, expected fragment size = 225bp) was
amplified from the genomic DNA of 100 healthy individuals. DQA was
amplified using standard PCR chemistry in a reaction having a total
volume of 50 μL containing of 10 mM Tris-HCl, 10 mM KCl (pH 8.3), 2.5
30   mM MgCl$_2$, 200 μM of each dNTP, 10 pmol of a forward primer having

the sequence 5'-GTG CTG CAG GTG TAA ACT TGT ACC AG-3'(SEQ ID NO: 64), 10 pmol of a reverse primer having the sequence 5'-CAC GGA TCC GGT AGC AGC GGT AGA GTT G-3'(SEQ ID NO: 65), 1 U DNA polymerase (Stoffel fragment, Perkin Elme r), and 200ng human genomic

5   DNA (2ng DNA/individual). The template was denatured at 94°C for 5 minutes. Thermal cycling was continued with a touch-down program that included 45 cycles of 20 seconds at 94°C, 30 seconds at 56°C, 1 minute at 72°C, and a final extension of 3 minutes at 72°C. The crude PCR product was used in the subsequent nickase reaction.

10   The unpurified PCR product was subjected to nickase digestion. Tris-HCl (10 mM), KCl (10 mM, pH 8.3), magnesium acetate (25mM), BSA (1 mg/mL), and 5 U of Cvi NY2A or Cvi NYS1 Nickase (Megabase Research) were added to 25 pmol of the amplified template with a total reaction volume of 20μL. The mixture was then incubated at 37°C for 5

15   hours. The digestion products were purified with either ZipTips (Millipore, Bedford, MA) as described in Example 5. The samples were analyzed by MALDI-TOF mass spectroscopy as described in Example 4. This assay can also be used to do multiplexing and standardless genotyping as described herein.

20   To simplify the nickase mass spectrum, the two complementary strands can be separated after digestion by using a single-stranded undigested PCR product as a capture probe. This probe (preparation shown below in Example 8C) can be hybridized to the nickase fragments in hybridization buffer containing 200 mM sodium citrate and 1% blocking

25   reagent (Boehringer Mannheim). The reaction is heated to 95°C for 5 minutes and cooled to room temperature over 30 minutes by using a thermal cycler (PTC-200 DNA engine, MJ Research, Waltham, MA). The capture probe-nickase fragment is immobilized on 140 μg of streptavidin-coated magnetic beads. The beads are subsequently washed three times

30   with 70 mM ammonium citrate. The captured single-stranded nickase

fragments are eluted by heating to 80°C for 5 minutes in 5 µL of 50 mM ammonium hydroxide.

C.   Preparation of Capture Probe

5   The capture probe is prepared by amplifying the human β-globin gene (3' end of intron 1 to 5' end of exon 2) via PCR methods in a total volume of 50 µL containing of GeneAmp 1XPCR Buffer II, 10 mM Tris-HCl, pH 8.3, 50 mM KCl, 2 mM MgCl$_2$, 0.2 mM dNTP mix, 10pmol of each primer (forward primer 5'-ACTGGGCATGTGGAGACAG-3'(SEQ ID NO: 66) and biotinylated reverse primer bio5'-GCACTTTCTTGCCATGAG-

10 3'(SEQ ID: 67), 2 U of AmpliTaq Gold, and 200 ng of human genomic DNA. The template is denatured at 94°C for 8 minutes. Thermal cycling is continued with a touch-down program that included 11 cycles of 20 seconds at 94°C, 30 seconds at 64°C, 1 minute at 72°C; and a final extension of 5 minutes at 72°C. The amplicon is purified using

15 UltraClean™ PCR clean-up kit (MO Bio Laboratories, Solano Beach, CA).

EXAMPLE 9

Multiplex Type IIS SNP Assay

A Type IIS assay was used to identify human gene sequences with

20 known SNPs. The Type IIS enzyme used in this assay was Fok I which effected double-stranded cleavage of the target DNA. The assay involved the steps of amplification and Fok I treatment of the amplicon. In the amplification step, the primers were designed so that each PCR product of a designated gene target was less than 100 bases such that a Fok I

25 recognition sequence was incorporated at the 5' and 3' end of the amplicon. Therefore, the fragments that were cleaved by Fok I included a center fragment containing the SNP of interest.

Ten human gene targets with known SNPs were analyzed by this assay. Sequences of the ten gene targets, as well as the primers used to

30 amplify the target regions, are found in Table 5. The ten targets were

-92-

lipoprotein lipase, prothrombin, factor V, cholesterol ester transfer protein (CETP), factor VII, factor XIII, HLA-H exon 2, HLA-H exon 4, methylenetetrahydrofolate reductase (MTHR), and P53 exon 4 codon 72.

5   Amplification of the ten human gene sequences were carried out in a single 50 μL volume PCR reaction with 20 ng of human genomic DNA template in 5 PCR reaction tubes. Each reaction vial contained 1X PCR buffer (Qiagen), 200μM dNTPs, 1U Hotstar Taq polymerase (Qiagen), 4 mM $MgCl_2$, and 10pmol of each primer. US8, having sequence of 5'TCAGTCACGACGTT3'(SEQ ID NO: 68), and US9, having sequence of

10   5'CGGATAACAATTTC3'(SEQ ID NO: 69), were used for the forward and reverse primers respectively. Moreover, the primers were designed such that a Fok I recognition site was incorporated at the 5' and 3' ends of the amplicon. Thermal cycling was performed in 0.2 mL tubes or a 96 well plate using a MJ Research Thermal Cycler (calculated temperature) with

15   the following cycling parameters: 94°C for 5 minutes; 45 cycles: 94°C for 20 seconds, 56°C for 20 seconds, 72°C for 60 seconds; and 72°C for 3 minutes.

Following PCR, the sample was treated with 0.2 U Exonuclease I (Amersham Pharmacia) and S Alkaline Phosphotase (Amersham

20   Pharmacia) to remove the unincorporated primers and dNTPs. Typically, 0.2 U of exonuclease I and SAP were added to 5 μL of the PCR sample. The sample was then incubated at 37°C for 15 minutes. Exonuclease I and SAP were then inactivated by heating the sample up to 85°C for 15 minutes. Fok I digestion was performed by adding 2 U of Fok I (New

25   England Biolab) to the 5 uL PCR sample and incubating at 37°C for 30 minutes. Since the Fok I restriction sites are located on both sides of the amplicon, the 5' and 3' cutoff fragments have higher masses than the center fragment containing the SNP. The sample was then purified by anion exchange and analyzed by MALDI-TOF mass spectrometry as

30   described in Example 4. The masses of the gene fragments from this

-93-

multiplexing experiment are listed in Table 6. These gene fragments were
resolved in mass spectra thereby allowing multiplex analysis of sequence
variability in these genes.

Table 5
Genes for Multiplex Type IIS Assay

Gene	Sequence	Seq. ID No.	Primers	Seq. ID No.
Lipoprotein Lipase (Asn291Ser)	cctttgagaa agggctctgc ttgagttgta gaaagaaccg ctgcaacaat ctggactatg agatca[a▸g]taa agtcagagcc aaaagaagca gcaaaatgta	98-99	5' caatttcatcgctggatgcaatct gggctatgagatc 3' 5' caatttcacacagcggatgcttct tttggctctgact 3'	70 71
Prothrombin	26731 gaattattttt tgtgtttcta aaactatggt tcccaataaa agtgactctc 26781 agc[g▸a]agcctc aatgctccca gtgctattca tgggcagctc tctgggctca	100-101	5' tcagtcacgacgttggatgccaa taaaagtgactctcagc 3' 5' cggataacaatttcggatgcact ggaagcattgaggc 3'	72 73
Factor V (Arg506Gln)	taataggact acttctaatc tgtaagagca gatcctgga caggc[g▸a]agga atacaggtat tttgtccttg aagtaacctt tcag	102-103	5' tcagtcacgacgttggatgagca gatcctgaacaggc 3' 5' cggataacaatttcggatggaca aaatacctgtattcc 3'	74 75
Cholesterol ester transfer protein (CETP) (I405V)	1261 ctcaccatgg gcatttgatt gcagagcagc tccgagtc[g▸a] tccagagctt 1311 cctgcagtca atgatcaccg ctgtgggcat ccctgaggtc atgtctgta	104-105	5' tcagtcacgacgttggatgcaga gcagcctccgagtc 3' 5' cagccggtgatcattggatgcagg aagctctgg 3'	76 77
Factor VII (R353Q)	1221 agcaaggact cctgcaaggg ggacagtgga ggccacatg ccacccacta 1271 cc[a▸g]gggcacg tggtacctga cgggcatcgt cagctggggc cagggctgcg	106-107	5' tcagtcacgacgttggatgccca catgccacccactac 3' 5' cggataacaatttcggatgcccg tcaggtaccacg 3'	78 79
Factor XIII (V34L)	111 caataatct aatgcagcgg aagatgacct gcccacagtg gagcttcagg 161 gc[g▸t]tggtgcc ccgggcgtc aacctgcaag gtatgagcat accccccttc	108-109	5' tcagtcacgacgttggatgccca cagtggagcttcag 3' 5' gctcataccttgcaggatgacg 3'	80 81
HLA-H exon 2 (His63Asp)	361 ttgaagcttt gggctacgtg gatgaccagc tgttcgtgtt ctatgat[c▸g]at 411 gagagtcgcc gtgtggagcc ccgaactcca tgggtttcca gtagaatttc	110-111	5' tcagtcacgacgttggatgcaga gctgttcgtgttc 3' 5' tacatggagttcgggggatgcaca cggcgactctc 3'	82 83

-94-

Gene	Sequence	Seq. ID No.	Primers	Seq. ID No.
HLA-H exon 4 (Cys282Tyr)	1021 ggataacctt ggctgtaccc cct<u>ggggaag agcaagagata</u> tacgt[g►a]ccag **1071 <u>gtggagcacc caggcctgga tcagcccctc</u>** attgtgatct gggagccctc	112-113	5' tcagtcacgacgttgga<u>tggggaa agagcagagatatacgt</u> 3' 5' <u>gagggggctgatccaggatgggt gctccac</u> 3'	84 85
Methylenetetrahy drofolateredctas e (MTHR) (Ala222Val)	761 <u>tgaagcactt gaagga gaag gtgtctgcgg</u> gag[c►t]<u>cgattt catcatcacg</u> 811 <u>cagcttttct</u> ttgaggctga cacattcttc	114-115	5' tcagtcacgacgttgga<u>tggggaa agagcagagatatacgt</u> 3' 5' <u>gagggggctgatccaggatgggt gctccac</u> 3'	86 87
P53 Exon4 Codon 72 (Arg72Pro)	1210 1 tccag<u>atgaa gctcccagaa tgccagagc</u> tgctcccc[g►c]c gtggcccctg 1215 1 <u>caccagcagc tcctacaccg gcggcccctg</u>	116-117	5' <u>gatgaagctcccaggatgccag aggc</u> 3' 5' <u>gccgccggtgtaggatgctgctg gtgc</u> 3'	88 89

Table 6
The mass of Center Fragments for Ten Different SNP Typing by IIS Assay

Gene	LPL(^{Asn}291^{Ser})		Prothrombin		FV(^{Arg}506^{Gln})		CETP(^{I}405^{V})		FVII(^{R}353^{Q})		FXIII(^{V}34)	
Genotype	A	G	G	A	G	A	G	A	G	A	G	T
+ strand mass (Da)	6213	6229	5845	5829	5677	5661	3388	3372	6128	6112	5058	5033
- strand mass (Da)	6129	6114	5949	5964	5472	5487	3437	3452	6174	6189	4916	4940

Gene	Hlah2		Hlah4		MTHR(^{Ala}222^{Val})		P53exon4(^{Arg}72^{Pro})	
Genotype	C	G	G	A	C	T	G	C
+ strand mass (Da)	5889	5929	4392	4376	4400	4415	4586	4546
- strand mass (Da)	5836	5796	4319	4334	4368	4352	4724	4764

EXAMPLE 10

Exemplary use of parental medical history parameter for stratification of healthy datebase

A healthy database can be used to associate a disease state with a specific allele (SNP) that has been found to show a strong association between age and the allele, in particular the homozygous genotype. The method involves using the same healthy database used to identify the age dependent association, however stratification is by information given by the donors about common disorders from which their parents suffered (the donor's familial history of disease). There are three possible answers a donor could give about the health status of their parents: neither were affected, one was affected or both were affected. Only donors above a certain minimum age, depending on the disease, are utilized, as the donors parents must be old enough to to have exhibited clinical disease phenotypes. The genotype frequency in each of these groups is determined and compared with each other. If there is an association of the marker in the donor to a disease the frequency of the heterozyous genotype will be increased. The frequency of the homozygous genotype should not increase, as it should be significantly underrepresented in the healthy population.

EXAMPLE 11

Method and Device for Identifying a Biological Sample

Description

In accordance with the present invention, a method and device for
5   identifying a biological sample is provided. Referring now to FIG. 24, an apparatus 10 for identifying a biological sample is disclosed. The apparatus 10 for identifying a biological sample generally comprises a mass spectrometer 15 communicating with a computing device 20. In a preferred embodiment, the mass spectrometer may be a MALDI-TOF mass spectrometer manufactured by
10   Bruker-Franzen Analytik GmbH; however, it will be appreciated that other mass spectrometers can be substituted. The computing device 20 is preferably a general purpose computing device. However, it will be appreciated that the computing device could be alternatively configured, for example, it may be

integrated with the mass spectrometer or could be part of a computer in a larger network system.

The apparatus 10 for identifying a biological sample may operate as an automated identification system having a robot 25 with a robotic arm 27
5   configured to deliver a sample plate 29 into a receiving area 31 of the mass spectrometer 15. In such a manner, the sample to be identified may be placed on the plate 29 and automatically received into the mass spectrometer 15. The biological sample is then processed in the mass spectrometer to generate data indicative of the mass of DNA fragments in the biological sample. This data may
10   be sent directly to computing device 20, or may have some preprocessing or filtering performed within the mass spectrometer. In a preferred embodiment, the mass spectrometer 15 transmits unprocessed and unfiltered mass spectrometry data to the computing device 20. However, it will be appreciated that the analysis in the computing device may be adjusted to accommodate
15   preprocessing or filtering performed within the mass spectrometer.

Referring now to FIG. 25, a general method 35 for identifying a biological sample is shown. In method 35, data is received into a computing device from a test instrument in block 40. Preferably the data is received in a raw, unprocessed and unfiltered form, but alternatively may have some form of
20   filtering or processing applied. The test instrument of a preferred embodiment is a mass spectrometer as described above. However, it will be appreciated that other test instruments could be substituted for the mass spectrometer.

The data generated by the test instrument, and in particular the mass spectrometer, includes information indicative of the identification of the
25   biological sample. More specifically, the data is indicative of the DNA composition of the biological sample. Typically, mass spectrometry data gathered from DNA samples obtained from DNA amplification techniques are noisier than, for example, those from typical protein samples. This is due in part because protein samples are more readily prepared in more abundance, and
30   protein samples are more easily ionizable as compared to DNA samples. Accordingly, conventional mass spectrometer data analysis techniques are generally ineffective for DNA analysis of a biological sample. To improve the

analysis capability so that DNA composition data can be more readily discerned, a preferred embodiment uses wavelet technology for analyzing the DNA mass spectrometry data. Wavelets are an analytical tool for signal processing, numerical analysis, and mathematical modeling. Wavelet technology provides a

5 basic expansion function which is applied to a data set. Using wavelet decomposition, the data set can be simultaneously analyzed in the time and frequency domains. Wavelet transformation is the technique of choice in the analysis of data that exhibit complicated time (mass) and frequency domain information, such as MALDI-TOF DNA data. Wavelet transforms as described

10 herein have superior denoising properties as compared to conventional Fourier analysis techniques. Wavelet transformation has proven to be particularly effective in interpreting the inherently noisy MALDI-TOF spectra of DNA samples. In using wavelets, a "small wave" or "scaling function" is used to transform a data set into stages, with each stage representing a frequency

15 component in the data set. Using wavelet transformation, mass spectrometry data can be processed, filtered, and analyzed with sufficient discrimination to be useful for identification of the DNA composition for a biological sample.

Referring again to FIG. 25, the data received in block 40 is denoised in block 45. The denoised data then has a baseline correction applied in block 50.

20 A baseline correction is generally necessary as data coming from the test instrument, in particular a mass spectrometer instrument, has data arranged in a generally exponentially decaying manner. This generally exponential decaying arrangement is not due to the composition of the biological sample, but is a result of the physical properties and characteristics of the test instrument, and

25 other chemicals involved in DNA sample preparation. Accordingly, baseline correction substantially corrects the data to remove a component of the data attributable to the test system, and sample preparation characteristics.

After denoising in block 45 and the baseline correction in block 50, a signal remains which is generally indicative of the composition of the biological

30 sample. However, due to the extraordinary discrimination required for analyzing the DNA composition of the biological sample, the composition is not readily apparent from the denoised and corrected signal. For example, although the

-98-

signal may include peak areas, it is not yet clear whether these "putative" peaks actually represent a DNA composition, or whether the putative peaks are result of a systemic or chemical aberration. Further, any call of the composition of the biological sample would have a probability of error which would be unacceptable for clinical or therapeutic purposes. In such critical situations, there needs to be a high degree of certainty that any call or identification of the sample is accurate. Therefore, additional data processing and interpretation is necessary before the sample can be accurately and confidently identified.

Since the quantity of data resulting from each mass spectrometry test is typically thousands of data points, and an automated system may be set to perform hundreds or even thousands of tests per hour, the quantity of mass spectrometry data generated is enormous. To facilitate efficient transmission and storage of the mass spectrometry data, block 55 shows that the denoised and baseline corrected data is compressed.

In a preferred embodiment, the biological sample is selected and processed to have only a limited range of possible compositions. Accordingly, it is therefore known where peaks indicating composition should be located, if present. Taking advantage of knowing the location of these expected peaks, in block 60 the method 35 matches putative peaks in the processed signal to the location of the expected peaks. In such a manner, the probability of each putative peak in the data being an actual peak indicative of the composition of the biological sample can be determined. Once the probability of each peak is determined in block 60, then in block 65 the method 35 statistically determines the composition of the biological sample, and determines if confidence is high enough to calling a genotype.

Referring again to block 40, data is received from the test instrument, which is preferably a mass spectrometer. In a specific illustration, FIG. 26 shows an example of data from a mass spectrometer. The mass spectrometer data 70 generally comprises data points distributed along an x-axis 71 and a y-axis 72. The x-axis 71 represents the mass of particles detected, while the y-axis 72 represents a numerical concentration of the particles. As can be seen in FIG. 26, the mass spectrometry data 70 is generally exponentially decaying with

-99-

data at the left end of the x-axis 73 generally decaying in an exponential manner toward data at the heavier end 74 of the x-axis 71. However, the general exponential presentation of the data is not indicative of the composition of the biological sample, but is more reflective of systematic error and characteristics.

5 Further, as described above and illustrated in FIG. 26, considerable noise exists in the mass spectrometry DNA data 70.

Referring again to block 45, where the raw data received in block 40 is denoised, the denoising process will be described in more detail. As illustrated in FIG. 25, the denoising process generally entails 1) performing a wavelet

10 transformation on the raw data to decompose the raw data into wavelet stage coefficients; 2) generating a noise profile from the highest stage of wavelet coefficients; and 3) applying a scaled noise profile to other stages in the wavelet transformation. Each step of the denoising process is further described below.

Referring now to FIG. 27, the wavelet transformation of the raw mass

15 spectrometry data is generally diagramed. Using wavelet transformation techniques, the mass spectrometry data 70 is sequentially transformed into stages. In each stage the data is represented in a high stage and a low stage, with the low stage acting as the input to the next sequential stage. For example, the mass spectrometry data 70 is transformed into stage 0 high data

20 82 and stage 0 low data 83. The stage 0 low data 83 is then used as an input to the next level transformation to generate stage 1 high data 84 and stage 1 low data 85. In a similar manner, the stage 1 low data 85 is used as an input to be transformed into stage 2 high data 86 and stage 2 low data 87. The transformation is continued until no more useful information can be derived by

25 further wavelet transformation. For example, in the preferred embodiment a 24-point wavelet is used. More particularly a wavelet commonly referred to as the Daubechies 24 is used to decompose the raw data. However, it will be appreciated that other wavelets can be used for the wavelet transformation. Since each stage in a wavelet transformation has one-half the data points of the

30 previous stage, the wavelet transformation can be continued until the stage n low data 89 has around 50 points. Accordingly, the stage n high 88 would contain about 100 data points. Since the preferred wavelet is 24 points long,

little data or information can be derived by continuing the wavelet transformation on a data set of around 50 points.

FIG. 28 shows an example of stage 0 high data 95. Since stage 0 high data 95 is generally indicative of the highest frequencies in the mass spectrometry data, stage 0 high data 95 will closely relate to the quantity of high frequency noise in the mass spectrometry data. In FIG. 29, an exponential fitting formula has been applied to the stage 0 high data 95 to generate a stage 0 noise profile 97. In particular, the exponential fitting formula is in the format $A_0 + A_1 EXP (-A_2 m)$. It will be appreciated that other expediential fitting formulas or other types of curve fits may be used.

Referring now to FIG. 30, noise profiles for the other high stages are determined. Since the later data points in each stage will likely be representative of the level of noise in each stage, only the later data points in each stage are used to generate a standard deviation figure that is representative of the noise content in that particular stage. More particularly, in generating the noise profile for each remaining stage, only the last five percent of the data points in each stage are analyzed to determined a standard deviation number. It will be appreciated that other numbers of points, or alternative methods could be used to generate such a standard deviation figure.

The standard deviation number for each stage is used with the stage 0 noise profile (the exponential curve) 97 to generate a scaled noise profile for each stage. For example, FIG. 30 shows that stage 1 high data 98 has stage 1 high data 103 with the last five percent of the data points represented by area 99. The points in area 99 are evaluated to determine a standard deviation number indicative of the noise content in stage 1 high data 103. The standard deviation number is then used with the stage 0 noise profile 97 to generate a stage 1 noise profile.

In a similar manner, stage 2 high 100 has stage 2 high data 104 with the last five percent of points represented by area 101. The data points in area 101 are then used to calculate a standard deviation number which is then used to scale the stage 0 noise profile 97 to generate a noise profile for stage 2 data. This same process is continued for each of the stage high data as shown by the

-101-

stage n high 105. For stage n high 105, stage n high data 108 has the last five percent of data points indicated in area 106. The data points in area 106 are used to determine a standard deviation number for stage n. The stage n standard deviation number is then used with the stage 0 noise profile 97 to

5 generate a noise profile for stage n. Accordingly, each of the high data stages has a noise profile.

FIG. 31 shows how the noise profile is applied to the data in each stage. Generally, the noise profile is used to generate a threshold which is applied to the data in each stage. Since the noise profile is already scaled to adjust for the

10 noise content of each stage, calculating a threshold permits further adjustment to tune the quantity of noise removed. Wavelet coefficients below the threshold are ignored while those above the threshold are retained. Accordingly, the remaining data has a substantial portion of the noise content removed.

Due to the characteristics of wavelet transformation, the lower stages,

15 such as stage 0 and 1, will have more noise content than the later stages such as stage 2 or stage n. Indeed, stage n low data is likely to have little noise at all. Therefore, in a preferred embodiment the noise profiles are applied more aggressively in the lower stages and less aggressively in the later stages. For example, FIG. 31 shows that stage 0 high threshold is determined by multiplying

20 the stage 0 noise profile by a factor of four. In such a manner, significant numbers of data points in stage 0 high data 95 will be below the threshold and therefore eliminated. Stage 1 high threshold 112 is set at two times the noise profile for the stage 1 high data, and stage 2 high threshold 114 is set equal to the noise profile for stage 2 high. Following this geometric progression, stage n

25 high threshold 116 is therefore determined by scaling the noise profile for each respective stage n high by a factor equal to $(1/2^{n-2})$. It will be appreciated that other factors may be applied to scale the noise profile for each stage. For example, the noise profile may be scaled more or less aggressively to accommodate specific systemic characteristics or sample compositions. As

30 indicated above, stage n low data does not have a noise profile applied as stage n low data 118 is assumed to have little or no noise content. After the scaled noise profiles have been applied to each high data stage, the mass spectrometry

-102-

data 70 has been denoised and is ready for further processing. A wavelet transformation of the denoised signal results in the sparse data set 120 as shown in FIG. 31.

Referring again to FIG. 25, the mass spectrometry data received in block
5   40 has been denoised in block 45 and is now passed to block 50 for baseline correction. Before performing baseline correction, the artifacts introduced by the wavelet transformation procedure are preferably removed. Wavelet transformation results vary slightly depending upon which point of the wavelet is used as a starting point. For example, the preferred embodiment uses the 24-
10   point Daubechies-24 wavelet. By starting the transformation at the 0 point of the wavelet, a slightly different result will be obtained than if starting at points 1 or 2 of the wavelet. Therefore, the denoised data is transformed using every available possible starting point, with the results averaged to determine a final denoised and shifted signal. For example, FIG. 33 shows that the wavelet
15   coefficient is applied 24 different times and then the results averaged to generate the final data set. It will be appreciated that other techniques may be used to accommodate the slight error introduced due to wavelet shifting.

The formula 125 is generally indicated in FIG. 33. Once the signal has been denoised and shifted, a denoised and shifted signal 130 is generated as
20   shown in FIG. 58. FIG. 34 shows an example of the wavelet coefficient 135 data set from the denoised and shifted signal 130.

FIG. 36 shows that putative peak areas 145, 147, and 149 are located in the denoised and shifted signal 150. The putative peak areas are systematically identified by taking a moving average along the signal 150 and identifying
25   sections of the signal 150 which exceed a threshold related to the moving average. It will be appreciated that other methods can be used to identify putative peak areas in the signal 150.

Putative peak areas 145, 147 and 149 are removed from the signal 150 to create a peak-free signal 155 as shown in FIG. 37. The peak-free signal 155
30   is further analyzed to identify remaining minimum values 157, and the remaining minimum values 157 are connected to generate the peak-free signal 155.

FIG. 38 shows a process of using the peak-free signal 155 to generate a baseline 170 as shown in FIG. 39. As shown in block 162, a wavelet transformation is performed on the peak-free signal 155. All the stages from the wavelet transformation are eliminated in block 164 except for the n low stage.

5 The n low stage will generally indicate the lowest frequency component of the peak-free signal 155 and therefore will generally indicate the system exponential characteristics. Block 166 shows that a signal is reconstructed from the n low coefficients and the baseline signal 170 is generated in block 168.

FIG. 39 shows a denoised and shifted data signal 172 positioned adjacent

10 a correction baseline 170. The baseline correction 170 is subtracted from the denoised and shifted signal 172 to generate a signal 175 having a baseline correction applied as shown in FIG. 40. Although such a denoised, shifted, and corrected signal is sufficient for most identification purposes, the putative peaks in signal 175 are not identifiable with sufficient accuracy or confidence to call

15 the DNA composition of a biological sample.

Referring again to FIG. 25, the data from the baseline correction 50 is now compressed in block 55, the compression technique used in a preferred embodiment is detailed in FIG. 41. In FIG. 41 the data in the baseline corrected data is presented in an array format 182 with x-axis points 183 having an

20 associated data value 184. The x-axis is indexed by the non-zero wavelet coefficients, and the associated value is the value of the wavelet coefficient. In the illustrated data example in table 182, the maximum value 184 is indicated to be 1000. Although a particularly advantageous compression technique for mass spectrometry data is shown, it will be appreciated that other compression

25 techniques can be used. Although not preferred, the data may also be stored without compression.

In compressing the data according to a preferred embodiment, an intermediate format 186 is generated. The intermediate format 186 generally comprises a real number having a whole number portion 188 and a decimal

30 portion 190. The whole number portion is the x-axis point 183 while the decimal portion is the value data 184 divided by the maximum data value. For

-104-

example, in the data 182 a data value "25" is indicated at x-axis point "100". The intermediate value for this data point would be "100.025".

From the intermediate compressed data 186 the final compressed data 195 is generated. The first point of the intermediate data file becomes the

5   starting point for the compressed data. Thereafter each data point in the compressed data 195 is calculated as follows: the whole number portion (left of the decimal) is replaced by the difference between the current and the last whole number. The remainder (right of the decimal) remains intact. For example, the starting point of the compressed data 195 is shown to be the same as the

10   intermediate data point which is "100.025". The comparison between the first intermediate data point "100.025" and the second intermediate data point "150.220" is "50.220". Therefore, "50.220" becomes the second point of the compressed data 195. In a similar manner, the second intermediate point is "150.220" and the third intermediate data point is "500.0001". Therefore, the

15   third compressed data becomes "350.000". The calculation for determining compressed data points is continued until the entire array of data points is converted to a single array of real numbers.

FIG. 42 generally describes the method of compressing mass spectrometry data, showing that the data file in block 201 is presented as an

20   array of coefficients in block 202. The data starting point and maximum is determined as shown in block 203, and the intermediate real numbers are calculated in block 204 as described above. With the intermediate data points generated, the compressed data is generated in block 205. The described compression method is highly advantageous and efficient for compressing data

25   sets such as a processed data set from a mass spectrometry instrument. The method is particularly useful for data, such as mass spectrometry data, that uses large numbers and has been processed to have occasional lengthy gaps in x-axis data. Accordingly, an x-y data array for processed mass spectrometry data may be stored with an effective compression rate of 10x or more. Although the

30   compression technique is applied to mass spectrometry data, it will be appreciated that the method may also advantageously be applied to other data sets.

Referring again to FIG. 25, peak heights are now determined in block 60. The first step in determining peak height is illustrated in FIG. 43 where the signal 210 is shifted left or right to correspond with the position of expected peaks. As the set of possible compositions in the biological sample is known before the

5   mass spectrometry data is generated, the possible positioning of expected peaks is already known. These possible peaks are referred to as expected peaks, such as expected peaks 212, 214, and 216. Due to calibration or other errors in the test instrument data, the entire signal may be shifted left or right from its actual position, therefore, putative peaks located in the signal, such as putative peaks

10 218, 222, and 224 may be compared to the expected peaks 212, 214, and 216, respectively. The entire signal is then shifted such that the putative peaks align more closely with the expected peaks.

Once the putative peaks have been shifted to match expected peaks, the strongest putative peak is identified in FIG. 44. In a preferred embodiment, the

15 strongest peak is calculated as a combination of analyzing the overall peak height and area beneath the peak. For example, a moderately high but wide peak would be stronger than a very high peak that is extremely narrow. With the strongest putative peak identified, such as putative peak 225, a Gaussian 228 curve is fit to the peak 225. Once the Gaussian is fit, the width (W) of the

20 Gaussian is determined and will be used as the peak width for future calculations.

As generally addressed above, the denoised, shifted, and baseline-corrected signal is not sufficiently processed for confidently calling the DNA composition of the biological sample. For example, although the baseline has

25 generally been removed, there are still residual baseline effects present. These residual baseline effects are therefore removed to increase the accuracy and confidence in making identifications.

To remove the residual baseline effects, FIG. 45 shows that the putative peaks 218, 222, and 224 are removed from the baseline corrected signal. The

30 peaks are removed by identifying a center line 230, 232, and 234 of the putative peaks 218, 222, and 224, respectively and removing an area to the left and to the right of the identified center line. For each putative peak, an area

equal to twice the width (W) of the Gaussian is removed from the left of the center line, while an area equivalent to 50 daltons is removed from the right of the center line. It has been found that the area representing 50 daltons is adequate to sufficiently remove the effect of salt adducts which may be

5   associated with an actual peak. Such adducts appear to the right of an actual peak and are a natural effect from the chemistry involved in acquiring a mass spectrum. Although a 50 Dalton buffer has been selected, it will be appreciated that other ranges or methods can be used to reduce or eliminate adduct effects.

     The peaks are removed and remaining minima 247 located as shown in

10 FIG. 46 with the minima 247 connected to create signal 245. A quartic polynomial is applied to signal 245 to generate a residual baseline 250 as shown in FIG. 47. The residual baseline 250 is subtracted from the signal 225 to generate the final signal 255 as indicated in FIG. 48. Although the residual baseline is the result of a quartic fit to signal 245, it will be appreciated that

15 other techniques can be used to smooth or fit the residual baseline.

     To determine peak height, as shown in FIG. 49, a Gaussian such as Gaussian 266, 268, and 270 is fit to each of the peaks, such as peaks 260, 262, and 264, respectively. Accordingly, the height of the Gaussian is determined as height 272, 274, and 276. Once the height of each Gaussian

20 peak is determined, then the method of identifying a biological compound 35 can move into the genotyping phase 65 as shown in FIG. 25.

     An indication of the confidence that each putative peak is an actual peak can be discerned by calculating a signal-to-noise ratio for each putative peak. Accordingly, putative peaks with a strong signal-to-noise ratio are generally more

25 likely to be an actual peak than a putative peak with a lower signal-to-noise ratio. As described above and shown in FIG. 50, the height of each peak, such as height 272, 274, and 276, is determined for each peak, with the height being an indicator of signal strength for each peak. The noise profile, such as noise profile 97, is extrapolated into noise profile 280 across the identified peaks. At

30 the center line of each of the peaks, a noise value is determined, such as noise value 282, 283, and 284. With a signal values and a noise values generated, signal-to-noise ratios can be calculated for each peak. For example, the signal-

to-noise ratio for the first peak in FIG. 50 would be calculated as signal value 272 divided by noise value 282, and in a similar manner the signal-to-noise ratio of the middle peak in FIG. 50 would be determined as signal 274 divided by noise value 283.

5   Although the signal-to-noise ratio is generally a useful indicator of the presence of an actual peak, further processing has been found to increase the confidence by which a sample can be identified. For example, the signal-to-noise ratio for each peak in the preferred embodiment is preferably adjusted by the goodness of fit between a Gaussian and each putative peak. It is a

10   characteristic of a mass spectrometer that sample material is detected in a manner that generally complies with a normal distribution. Accordingly, greater confidence will be associated with a putative signal having a Gaussian shape than a signal that has a less normal distribution. The error resulting from having a non-Gaussian shape can be referred to as a "residual error".

15   Referring to FIG. 51, a residual error is calculated by taking a root mean square calculation between the Gaussian 293 and the putative peak 290 in the data signal. The calculation is performed on data within one width on either side of a center line of the Gaussian. The residual error is calculated as:

       where G is the Gaussian signal value, R is the putative peak value, and N

20   is the number of points from $-W$ to $+W$. The calculated residual error is used to generate an adjusted signal-to-noise ratio, as described below.

       An adjusted signal noise ratio is calculated for each putative peak using the formula $(S/N) * EXP^{(-1 * R)}$, where S/N is the signal-to-noise ratio, and R is the residual error determined above. Although the preferred embodiment

25   calculates an adjusted signal-to-noise ratio using a residual error for each peak, it will be appreciated that other techniques can be used to account for the goodness of fit between the Gaussian and the actual signal.

       Referring now to FIG. 52, a probability is determined that a putative peak is an actual peak. In making the determination of peak probability, a probability

30   profile 300 is generated where the adjusted signal-to-noise ratio is the x-axis and the probability is the y-axis. Probability is necessarily in the range between a 0% probability and a 100% probability, which is indicated as 1. Generally, the

-108-

higher the adjusted signal-to-noise ratio, the greater the confidence that a putative peak is an actual peak.

At some target value for the adjusted signal-to-noise, it has been found that the probability is 100% that the putative peak is an actual peak and can
5   confidently be used to identify the DNA composition of a biological sample. However, the target value of adjusted signal-to-noise ratio where the probability is assumed to be 100% is a variable parameter which is to be set according to application specific criteria. For example, the target signal-to-noise ratio will be adjusted depending upon trial experience, sample characteristics, and the
10 acceptable error tolerance in the overall system. More specifically, for situations requiring a conservative approach where error cannot be tolerated, the target adjusted signal-to-noise ratio can be set to, for example, 10 and higher. Accordingly, 100% probability will not be assigned to a peak unless the adjusted signal-to-noise ratio is 10 or over.

15 In other situations, a more aggressive approach may be taken as sample data is more pronounced or the risk of error may be reduced. In such a situation, the system may be set to assume a 100% probability with a 5 or greater target signal-to-noise ratio. Of course, an intermediate signal-to-noise ratio target figure can be selected, such as 7, when a moderate risk of error can
20 be assumed. Once the target adjusted signal-to-noise ratio is set for the method, then for any adjusted signal-to-noise ratio a probability can be determined that a putative peak is an actual peak.

Due to the chemistry involved in performing an identification test, especially a mass spectrometry test of a sample prepared by DNA amplifications,
25 the allelic ratio between the signal strength of the highest peak and the signal strength of the second (or third and so on) highest peak should fall within an expected ratio. If the allelic ratio falls outside of normal guidelines, the preferred embodiment imposes an allelic ratio penalty to the probability. For example, FIG. 53 shows an allelic penalty 315 which has an x-axis 317 that is the ratio
30 between the signal strength of the second highest peak divided by signal strength of the highest peak. The y-axis 319 assigns a penalty between 0 and 1 depending on the determined allelic ratio. In the preferred embodiment, it is

-109-

assumed that allelic ratios over 30% are within the expected range and therefore no penalty is applied. Between a ratio of 10% and 30%, the penalty is linearly increased until at allelic ratios below 10% it is assumed the second-highest peak is not real. For allelic ratios between 10% and 30%, the allelic penalty chart

5   315 is used to determine a penalty 319, which is multiplied by the peak probability determined in FIG. 52 to determine a final peak probability. Although the preferred embodiment incorporates an allelic ratio penalty to account for a possible chemistry error, it will be appreciated that other techniques may be used. Similar treatment will be applied to the other peaks.

10 With the peak probability of each peak determined, the statistical probability for various composition components may be determined. As an example, in order to determine the probability of each of three possible combinations of two peaks, -- peak G, peak C and combinations GG, CC and GC. FIG. 54 shows an example where a most probable peak 325 is determined

15 to have a final peak probability of 90%. Peak 325 is positioned such that it represents a G component in the biological sample. Accordingly, it can be maintained that there is a 90% probability that G exists in the biological sample. Also in the example shown in FIG. 54, the second highest probability is peak 330 which has a peak probability of 20%. Peak 330 is at a position associated

20 with a C composition. Accordingly, it can be maintained that there is a 20% probability that C exists in the biological sample.

With the probability of G existing (90%) and the probability of C existing (20%) as a starting point, the probability of combinations of G and C existing can be calculated. For example, FIG. 54 indicates that the probability of GG

25 existing 329 is calculated as 72%. This is calculated as the probability of GG is equal to the probability of G existing (90%) multiplied by the probability of C not existing (100% -20%). So if the probability of G existing is 90% and the probability of C not existing is 80%, the probability of GG is 72%.

In a similar manner, the probability of CC existing is equivalent to the

30 probability of C existing (20%) multiplied by the probability of G not existing (100% - 90%). As shown in FIG. 54, the probability of C existing is 20% while the probability of G not existing is 10%, so therefore the probability of CC is

only 2%. Finally, the probability of GC existing is equal to the probability of G existing (90%) multiplied by the probability of C existing (20%). So if the probability of G existing is 90% and the probability of C existing is 20%, the probability of GC existing is 18%. In summary form, then, the probability of the composition of the biological sample is:

probability of GG: 72%;

probability of GC: 18%; and

probability of CC: 2%.

Once the probabilities of each of the possible combinations has been determined, FIG. 55 is used to decide whether or not sufficient confidence exists to call the genotype. FIG. 55 shows a call chart 335 which has an x-axis 337 which is the ratio of the highest combination probability to the second highest combination probability. The y-axis 339 simply indicates whether the ratio is sufficiently high to justify calling the genotype. The value of the ratio may be indicated by M 340. The value of M is set depending upon trial data, sample composition, and the ability to accept error. For example, the value M may be set relatively high, such as to a value 4 so that the highest probability must be at least four times greater than the second highest probability before confidence is established to call a genotype. However, if a certain level of error may be acceptable, the value of M may be set to a more aggressive value, such as to 3, so that the ratio between the highest and second highest probabilities needs to be only a ratio of 3 or higher. Of course, moderate value may be selected for M when a moderate risk can be accepted. Using the example of FIG. 54, where the probability of GG was 72% and the probability of GC was 18%, the ratio between 72% and 18% is 4.0, therefore, whether M is set to 3, 3.5, or 4, the system would call the genotype as GG. Although the preferred embodiment uses a ratio between the two highest peak probabilities to determine if a genotype confidently can be called, it will be appreciated that other methods may be substituted. It will also be appreciated that the above techniques may be used for calculating probabilities and choosing genotypes (or more general DNA patterns) containing of combinations of more than two peaks.

Referring now to FIG. 56, a flow chart is shown generally defining the process of statistically calling genotype described above. In FIG. 56 block 402 shows that the height of each peak is determined and that in block 404 a noise profile is extrapolated for each peak. The signal is determined from the height of each peak in block 406 and the noise for each peak is determined using the noise profile in block 408. In block 410, the signal-to-noise ratio is calculated for each peak. To account for a non-Gaussian peak shape, a residual error is determined in block 412 and an adjusted signal-to-noise ratio is calculated in block 414. Block 416 shows that a probability profile is developed, with the probability of each peak existing found in block 418. An allelic penalty may be applied in block 420, with the allelic penalty applied to the adjusted peak probability in block 422. The probability of each combination of components is calculated in block 424 with the ratio between the two highest probabilities being determined in block 426. If the ratio of probabilities exceeds a threshold value then the genotype is called in block 428.

In another embodiment of the invention, the computing device 20 (Fig. 24) supports "standardless" genotyping by identifying data peaks that contain putative SNPs. Standardless genotyping is used, for example, where insufficient information is known about the samples to determine a distribution of expected peak locations, against which an allelic penalty as described above can be reliably calculated. This permits the computing device to be used for identification of peaks that contain putative SNPs from data generated by any assay that fragments a targeted DNA molecule. For such standardless genotyping, peaks that are associated with an area under the data curve that deviates significantly from the typical area of other peaks in the data spectrum are identified and their corresponding mass (location along the x-axis) is determined.

More particularly, peaks that deviate significantly from the average area of other peaks in the data are identified, and the expected allelic ratio between data peaks is defined in terms of the ratio of the area under the data peaks. Theoretically, where each genetic loci has the same molar concentration of

-112-

analyte, the area under each corresponding peak should be the same, thus producing a 1.0 ratio of the peak area between any two peaks. In accordance with the invention, peaks having a smaller ratio relative to the other peaks in the data will not be recognized as peaks. More particularly, peaks having an area

5   ratio smaller than 30% relative to a nominal value for peak area will be assigned an allelic penalty. The mass of the remaining peaks (their location along the x-axis of the data) will be determined based on oligonucleotide standards.

Fig. 57 shows a flow diagram representation of the processing by the computing device 20 (Fig. 24) when performing standardless genotyping. In the

10   first operation, represented by the flow diagram box numbered 502, the computing device receives data from the mass spectrometer. Next, the height of each putative peak in the data sample is determined, as indicated by the block 504. After the height of each peak in the mass spectrometer data is determined, a de-noise process 505 is performed, beginning with an

15   extrapolation of the noise profile (block 506), followed by finding the noise of each peak (block 508) and calculating the signal to noise ratio for each data sample (block 510). Each of these operations may be performed in accordance with the description above for denoise operations 45 of Fig. 25. Other suitable denoise operations will occur to those skilled in the art.

20   The next operation is to find the residual error associated with each data point. This is represented by the block 512 in Figure 57. The next step, block 514, involves calculating an adjusted signal to noise ratio for each identified peak. A probability profile is developed next (block 516), followed by a determination of the peak probabilities at block 518. In the preferred

25   embodiment, the denoise operations of Fig. 57, comprising block 502 to block 518, comprise the corresponding operations described above in conjunction with Fig. 56 for block 402 through block 418, respectively.

The next action for the standardless genotype processing is to determine an allelic penalty for each peak, indicated by the block 524. As noted above,

30   the standardless genotype processing of Fig. 57 determines an allelic penalty by comparing area under the peaks. Therefore, rather than compare signal strength ratios to determine an allelic penalty, such as described above for Fig. 53, the

standardless processing determines the area under each of the identified peaks and compares the ratio of those areas. Determining the area under each peak may be computed using conventional numerical analysis techniques for calculating the area under a curve for experimental data.

5   Thus, the allelic penalty is assigned in accordance with Fig. 58, which shows that no penalty is assigned to peaks having a peak area relative to an expected average area value that is greater than 0.30 (30%). The allelic penalty is applied to the peak probability value, which may be determined according to the process such as described in Fig. 52. It should be apparent from Fig. 58

10   that the allelic penalty imposed for peaks below a ratio of 30% is that such peaks will be removed from further measurement and processing. Other penalty schemes, however, may be imposed in accordance with knowledge about the data being processed, as determined by those skilled in the art.

After the allelic penalty has been determined and applied, the

15   standardless genotype processing compares the location of the remaining putative peaks to oligonucleotide standards to determine corresponding masses in the processing for block 524. For standardless genotype data, the processing of the block 524 is performed to determine mass and genotype, rather than performing the operations corresponding to block 424, 426, and 428 of Fig. 33.

20   Techniques for performing such comparisons and determining mass will be known to those skilled in the art.

In another embodiment, the computing device 20 (Fig. 24) permits the detection and determination of the mass (location along the x-axis of the data) of the sense and antisense strand of fragments generated in the assay. If desired,

25   the computing device may also detect and determine the quantity (area under each peak) of the respective sense and antisense strands, using a similar technique to that described above for standardless genotype processing. The data generated for each type of strand may then be combined to achieve a data redundancy and to thereby increase the confidence level of the determined

30   genotype. This technique obviates primer peaks that are often observed in data from other diagnostic methods, thereby permitting a higher level of multiplexing. In addition, when quantitation is used in pooling experiments, the ratio of the

-114-

measured peak areas is more reliably calculated than the peak identifying technique, due to data redundancy.

Fig. 23 is a flow diagram that illustrates the processing implemented by the computing device 20 to perform sense and antisense processing. In the first operation, represented by the flow diagram box numbered 602, the computing device receives data from the mass spectrometer. This data will include data for the sense strand and antisense strand of assay fragments. Next, the height of each putative peak in the data sample is determined, as indicated by the block 604. After the height of each peak in the mass spectrometer data is determined, a de-noise process 605 is performed, beginning with an operation that extrapolates the noise profile (block 606), followed by finding the noise of each peak (block 608) and calculating the signal to noise ratio for each data sample (block 610). Each of these operations may be performed in accordance with the description above for the denoise operations 45 of Fig. 25. Other suitable denoise operations will occur to those skilled in the art. The next operation is to find the residual error associated with each data point. This is represented by the block 612 in Figure 36.

After the residual error for the data of the sense strand and antisense strand has been performed, processing to identify the genotypes will be performed for the sense strand and also for the antisense strand. Therefore, Fig. 23 shows that processing includes sense strand processing (block 630) and antisense strand processing (block 640). Each block 630, 640 includes processing that corresponds to adjusting the signal to noise ratio, developing a probability profile, determining an allelic penalty, adjusting the peak probability by the allelic penalty, calculating genotype probabilities, and testing genotype probability ratios, such as described above in conjunction with blocks 414 through 426 of Fig. 56. The processing of each block 630, 640 may, if desired, include standardless processing operations such as described above in conjunction with Fig. 57. The standardless processing may be included in place of or in addition to the processing operations of Fig. 56.

After the genotype probability processing is completed, the data from the sense strand and antisense strand processing is combined and compared to

-115-

expected database values to obtain the benefits of data redundancy as between the sense strand and antisense strand. Those skilled in the art will understand techniques to take advantage of known data redundancies between a sense strand and antisense strand of assay fragments. This processing is represented

5   by the block 650. After the data from the two strands is combined for processing, the genotype processing is performed (block 660) and the genotype is identified.

Since modifications will be apparent to those of skill in this art, it is
10   intended that this invention be limited only by the scope of the appended claims.

WHAT IS CLAIMED IS:

1.   A subcollection of samples from a target population, comprising:

a plurality of samples, wherein the samples are selected from the group consisting of blood, tissue, body fluid, cell, seed, microbe, pathogen and

5   reproductive tissue samples; and

a symbology on the containers containing the samples, wherein the symbology is representative of the source and/or history of each sample, wherein:

the target population is a healthy population that has not been selected

10   for any disease state;

the collection comprises samples from the healthy population; and

the subcollection is obtained by sorting the collection according to specified parameters.

2.   The subcollection of claim 1, wherein the parameters are selected

15   from the group consisting of ethnicity, age, gender, height, weight, alcohol intake, number of pregnancies, number of live births, vegetarians, type of physical activity, state of residence and/or length of residence in a particular state, educational level, age of parent at death, cause of parent death, former or current smoker, length of time as a smoker, frequency of smoking, occurrence

20   of a disease in immediate family (parent, siblings, children), use of prescription drugs and/or reason therefor, length and/or number of hospital stays and exposure to environmental factors.

3.   The subcollection of claim 1, wherein the symbology is a bar code.

4.   A method of producing a database, comprising:

25   identifying healthy members of a population;

obtaining data comprising identifying information and obtaining historical information and data relating to the identified members of the population and their immediate family;

entering the data into a database for each member of the population and

30   associating the member and the data with an indexer.

5.   The method of claim 4, further comprising:

obtaining a body tissue or body fluid sample;

-117-

analyzing the body tissue or body fluid in the sample; and

entering the results of the analysis for each member into the database and associating each result with the indexer representative of each member.

6. A database produced by the method of claim 4.

5   7. A database produced by the method of claim 5.

8. A database, comprising:

datapoints representative of a plurality of healthy organisms from whom biological samples are obtained,

wherein each datapoint is associated with data representative of

10   the organism type and other identifying information.

9.   The database of claim 8, wherein the datapoints are answers to questions regarding one or more of a parameters selected from the group consisting of ethnicity, age, gender, height, weight, alcohol intake, number of pregnancies, number of live births, vegetarians, type of physical activity, state of

15   residence and/or length of residence in a particular state, educational level, age of parent at death, cause of parent death, former or current smoker, length of time as a smoker, frequency of smoking, occurrence of a disease in immediate family (parent, siblings, children), use of prescription drugs and/or reason therefor, length and/or number of hospital stays and exposure to environmental

20   factors.

10.   The database of claim 9, wherein the organisms are mammals and the samples are body fluids or tissues.

11.   The database of claim 9, wherein the samples are selected from blood, blood fractions, cells and subcellular organelles.

25   12   The database of claim 8, further comprising,

phenotypic data from an organism.

13. The database of claim 12, wherein the data includes one of physical characteristics, background data, medical data, and historical data.

14.   The database of claim 8, further comprising,

30   genotypic data from nucleic acid obtained from an organism.

-118-

15. The database of claim 14, wherein genotypic data includes, genetic markers, non-coding regions, microsatellites, RFLPs, VNTRs, historical data of the organism, medical history, and phenotypic information.

16. The database of claim 8 that is a relational database.

5   17. The database of claim 16, wherein the data are related to an indexer datapoint representative of each organism from whom data is obtained.

18. A method of identifying polymorphisms that are candidate genetic markers, comprising:

identifying a polymorphism; and

10   identifying any pathway or gene linked to the locus of the polymorphism, wherein

the polymorphisms are identified in samples associated with a target population that comprises healthy subjects.

19. The method of claim 18, wherein the polymorphism is identified by

15 detecting the presence of target nucleic acids in a sample by a method, comprising the steps of:

a) hybridizing a first oligonucleotide to the target nucleic acid;

b) hybridizing a second oligonucleotide to an adjacent region of the target nucleic acid;

20   c) ligating the hybridized oligonucleotides; and

c) detecting hybridized first oligonucleotide by mass spectrometry as an indication of the presence of the target nucleic acid.

20. The method of claim 18, wherein the polymorphism is identified by detecting target nucleic acids in a sample by a method, comprising the steps of:

25   a) hybridizing a first oligonucleotide to the target nucleic acid and hybridizing a second oligonucleotide to an adjacent region of the target nucleic acid;

b) contacting the hybridized first and second oligonucleotides with a cleavage enzyme to form a cleavage product; and

30   c) detecting the cleavage product by mass spectrometry as an indication of the presence of the target nucleic acid.

21.   The method of claim 20 wherein the samples are from subjects in a healthy database.

22.   The method of claim 18, wherein the polymorphism is identified by identifying target nucleic acids in a sample by primer oligo base extension (probe).

23.   The method of 22, wherein primer oligo base extension, comprises:

a) obtaining a nucleic acid molecule that contains a target nucleotide;

b) optionally immobilizing the nucleic acid molecule onto a solid support, to produce an immobilized nucleic acid molecule;

c) hybridizing the nucleic acid molecule with a primer oligonucleotide that is complementary to the nucleic acid molecule at a site adjacent to the target nucleotide;

d) contacting the product of step c) with a composition comprising a dideoxynucleoside triphosphate or a 3'-deoxynucleoside triphosphates and a polymerase, so that only a dideoxynucleoside or 3'-deoxynucleoside triphosphate that is complementary to the target nucleotide is extended onto the primer; and

e) detecting the extended primer, thereby identifying the target nucleotide.

24.   The method of claim 23, wherein detection of the extended primer is effected by mass spectrometry, comprising:

ionizing and volatizing the product of step d) ; and

detecting the extended primer by mass spectrometry, thereby identifying the target nucleotide.

25.   The method of claim 24, wherein;

samples are presented to the mass spectrometer as arrays on chips; and

each sample occupies a volume that is about the size of the laser spot projected by the laser in a mass spectrometer used in matrix-assisted laser desorption/ionization (MALDI) spectrometry.

-120-

26.   A combination, comprising:

a database containing parameters associated with a datapoint representative of a subject from whom samples are obtained, wherein the subjects are healthy; and

5   an indexed collection of the samples, wherein the index identifies the subject from whom the sample was obtained.

27   The combination of claim 26, wherein the parameter is selected from the group consisting of ethnicity, age, gender, height, weight, alcohol intake, number of pregnancies, number of live births, vegetarians, type of

10   physical activity, state of residence and/or length of residence in a particular state, educational level, age of parent at death, cause of parent death, former or current smoker, length of time as a smoker, frequency of smoking, occurrence of disease in immediate family (parent, siblings, children), use of prescription drugs and/or reason therefor, length and/or number of hospital stays and

15   ecposure to environmental factors.

28.   The combination of claim 26, wherein the database further contains genotypic data for each subject.

29.   The combination of claim 26, wherein the samples are blood.

30   A data storage medium, comprising the database of claim 8.

20   31.   A computer system, comprising the database of claim 8.

32.   A system for high throughput processing of biological samples, comprising:

a process line comprising a plurality of processing stations, each of which
    performs a procedure on a biological sample contained in a
25   reaction vessel;

a robotic system that transports the reaction vessel from processing
    station to processing station;

a data analysis system that receives test results of the process line and
    automatically processes the test results to make a determination
30   regarding the biological sample in the reaction vessel;

a control system that determines when the test at each processing
    station is complete and, in response, moves the reaction vessel to

-121-

the next test station, and continuously processes reaction vessels one after another until the control system receives a stop instruction; and

a database of claim 8, wherein the samples tested by the automated process line comprise samples from subjects in the database.

33.   The system of claim 32, wherein one of the processing stations comprises a mass spectrometer.

34.   The system of claim 32, wherein the data analysis system processes the test results by receiving test data from the mass spectrometer such that the test data for a biological sample contains one or more signals, whereupon the data analysis system determines the area under the curve of each signal and normalizes the results thereof and obtains a substantially quantitative result representative of the relative amounts of components in the tested sample.

35.   A method for high throughput processing of biological samples, the method comprising:

transporting a reaction vessel along a system of claim 32, comprising a process line having a plurality of processing stations, each of which performs a procedure on one or more biological samples contained in the reaction vessel;

determining when the test procedure at each processing station is complete and, in response, moving the reaction vessel to the next processing station;

receiving test results of the process line and automatically processing the test results to make a data analysis determination regarding the biological samples in the reaction vessel; and

processing reaction vessels continuously one after another until receiving a stop instruction, wherein the samples tested by the automated process line comprise samples from subjects in the database.

36.   The method of 35, wherein one of the processing stations comprises a mass spectrometer.

-122-

37.   The method of claim 36, wherein the samples are analyzed by a method comprising primer oligo base extension (probe).

38.   The method of claim 37, further comprising:

processing the test results by receiving test data from the mass

5   spectrometer such that the test data for a biological sample contains one or more signals or numerical values representative of signals, whereupon the data analysis system determines the area under the curve of each signal and normalizes the results thereof and obtains a substantially quantitative result representative of the relative amounts of components in the tested sample.

10   39. The method of claim 37, wherein primer oligo base extension, comprises:

a) obtaining a nucleic acid molecule that contains a target nucleotide;

b) optionally immobilizing the nucleic acid molecule onto a solid support, to produce an immobilized nucleic acid molecule;

15   c) hybridizing the nucleic acid molecule with a primer oligonucleotide that is complementary to the nucleic acid molecule at a site adjacent to the target nucleotide;

d) contacting the product of step c) with composition comprising a dideoxynucleoside triphosphate or a 3'-deoxynucleoside triphosphates and a

20   polymerase, so that only a dideoxynucleoside or 3'-deoxynucleoside triphosphate that is complementary to the target nucleotide is extended onto the primer; and

e) detecting the primer, thereby identifying the target nucleotide.

40.   The method of 39, wherein detection of the extended primer is effected by mass spectrometry, comprising:

25   ionizing and volatizing the product of step d); and

detecting the extended primer by mass spectrometry, thereby identifying the target nucleotide.

41.   The method of claim 36, wherein the target nucleic acids in the sample are detected and/or identified by a method, comprising the steps of:

30   a)   hybridizing a first oligonucleotide to the target nucleic acid;

b)   hybridizing a second oligonucleotide to an adjacent region of the target nucleic acid;

-123-

c) ligating then hybridized oligonucleotides; and

c) detecting hybridized first oligonucleotide by mass spectrometry as an indication of the presence of the target nucleic acid.

42. The method of claim 36, wherein the target nucleic acids in the sample are detected and/or identified by a method, comprising the steps of:

a) hybridizing a first oligonucleotide to the target nucleic acid and hybridizing a second oligonucleotide to an adjacent region of the target nucleic acid;

b) contacting the hybridized first and second oligonucleotides with a cleavage enzyme to form a cleavage product; and

c) detecting the cleavage product by mass spectrometry as an indication of the presence of the target nucleic acid.

43. A method of producing a database stored in a computer memory, comprising:

identifying healthy members of a population;

obtaining identifying and historical information and data relating to the identified members of the population;

entering the member-related data into the computer memory database for each identified member of the population and associating the member and the data with an indexer.

44. The method of claim 43, further comprising:

obtaining a body tissue or body fluid sample of an identified member;

analyzing the body tissue or body fluid in the sample; and

entering the results of the analysis for each member into the computer memory database and associating each result with the indexer representative of each member.

45. A database produced by the method of claim 43.

46. A database produced by the method of claim 44.

47. The database of claim 8, wherein:

the organims are selected from among animals, bacteria, fungi, protozoans and parasites and

each datapoint is associated with parameters representative of the organism type and identifying information.

48. The database of claim 43, further comprising, phenotypic data regarding each subject.

5 49. The database of claim 47 that is a relational database and the parameters are the answers to the questions in the questionnaire.

50. The database of claim 8, further comprising,

genotypic data of nucleic acid of the subject, wherein genotypic data includes, but is not limited to, genetic markers, non-coding regions,

10 microsatellites, restriction fragment length polymorphisms (RFLPs), variable number tandem repeats (VNTRs), historical day of the organism, the medical history of the subject, phenotypic information, and other information.

51. A database, comprising data records stored in computer memory, wherein the data records contain information that identifies healthy members of

15 a population, and also contain identifying and historical information and data relating to the identified members.

52. The database of claim 51, further comprising an index value for each identified member that associates each member of the population with the identifying and historical information and data.

20 53. A computer system, comprising the database of claim 51.

54. An automated process line, comprising the database of claim 51.

55. A method for determining a polymorphism that correlates with age, ethnicity or gender, comprising:

identifying a polymorphism; and

25 determining the frequency of the polymorphism with increasing age, with ethnicity or with gender in a healthy population.

56. A method for determining whether a polymorphism correlates with suceptibility to morbidity, early mortality, or morbidity and early mortality, comprising;

30 identifying a polymorphism; and

determining the frequency of the polymorphism with increasing age in a healthy population.

-125-

57.   A high throughput method of determining frequencies of genetic variations, comprising:

selecting a healthy target population and a genetic variation to be assessed;

5   pooling a plurality of samples of biopolymers obtained from members of the population,

determining or detecting the biopolymer that comprises the variation by mass spectrometry;

obtaining a mass spectrum or a digital representation thereof; and

10   determining the frequency of the variation in the population.

58.   The method of claim 57, wherein:

the variation is selected from the group consisting of an allelic variation, a post-translational modification, a nucleic modification, a label, a mass modification of a nucleic acid and methylation; and/or

15   the biopolymer is a nucleic acid, a protein, a polysaccharide, a lipid, a small organic metabolite or intermediate, wherein the concentration of biopolymer of interest is the same in each of the samples; and/or

the frequency is determined by assessing the method comprising determining the area under the peak in the mass spectrum or digital

20   repesentation thereof corresponding to the mass of the biopolymer comprising the genomic variation.

59.   The method of claim 58, wherein the method for determining the frequency is effected by determining the ratio of the signal or the digital representation thereof to the total area of the entire mass spectrum, which is

25   corrected for background.

60.   A method for discovery of a polymorphism in a population, comprising:

sorting the database of claim 8 according to a selected parameter to identify samples that match the selected parameter;

30   isolating a biopolymer from each identified sample;

optionally pooling each isolated biopolymer;

optionally amplifying the amount of biopolymer;

-126-

cleaving the pooled biopolymers to produce fragments thereof;

obtaining a mass spectrum of the resulting fragments and comparing the mass spectrum with a control mass spectrum to identify differences between the spectra and thereby identifying any polymorphisms; wherein:

5   the control mass spectrum is obtained from unsorted samples in the collection or samples sorted according to a different parameter.

61.   The method of claim 60, wherein cleaving is effected by contacting the biopolymer with an enzyme.

62.   The method of claim 61, wherein the enzyme is selected from the 10   group consisting of nucleotide glycosylase, a nickase and a type IIS restriction enzyme.

63.   The method of claim 60, wherein the biopolymer is a nucleic acid or a protein.

64.   The method of claim 60, wherein the the mass spectrometric 15   format is selected from among Matrix-Assisted Laser Desorption/Ionization, Time-of-Flight (MALDI-TOF), Electrospray (ES), IR-MALDI, Ion Cyclotron Resonance (ICR), Fourier Transform and combinations thereof.

65.   A method for discovery of a polymorphism in a population, comprising:

20   obtaining samples of body tissue or fluid from a plurality of organisms;

isolating a biopolymer from each sample;

pooling each isolated biopolymer;

optionally amplifying the amount of biopolymer;

cleaving the pooled biopolymers to produce fragments thereof;

25   obtaining a mass spectrum of the resulting fragments;

comparing the frequency of each fragment to identify fragments present in amounts lower than the average frequency, thereby identifying any polymorphisms.

66. The method of claim 65, wherein cleaving is effected by contacting 30   the biopolymer with an enzyme.

-127-

67. The method of claim 66, wherein the enzyme is selected from the group consisting of nucleotide glycosylase, a nickase and a type IIS restriction enzyme.

68. The method of claim 65, wherein the biopolymer is a nucleic acid or a protein.

69. The method of claim 65, wherein the the mass spectrometric format is selected from among Matrix-Assisted Laser Desorption/Ionization, Time-of-Flight (MALDI-TOF), Electrospray (ES), IR-MALDI, Ion Cyclotron Resonance (ICR), Fourier Transform and combinations thereof.

70. The method of claim 65, wherein the samples are obtained from healthy subjects.

71. A method of correlating a polymorphism with a parameter, comprising:

sorting the database of claim 8 according to a selected parameter to identify samples that match the selected parameter;

isolating a biopolymer from each identified sample;

pooling each isolated biopolymer;

optionally amplifying the amount of biopolymer;

determining the frequency of the polymorphism in the pooled biopolymers, wherein:

an alteration of the frequency of the polymorphism compared to a control, indicates a correlation of the polymorphism with the selected parameter; and

the control is the frequency of the polymorphism in pooled biopolymers obtained from samples identified from an unsorted database or from a database sorting according to a different parameter.

72. The method claim 71, wherein the parameter is selected from the group consisting of ethnicity, age, gender, height, weight, alcohol intake, number of pregnancies, number of live births, vegetarians, type of physical activity, state of residence and/or length of residence in a particular state, educational level, age of parent at death, cause of parent death, former or current smoker, length of time as a smoker, frequency of smoking, occurrence of a disease in immediate family (parent, siblings, children), use of prescription

-128-

drugs and/or reason therefor, length and/or number of hospital stays and exposure to environmental factors.

       73.   The method claim 72, wherein the parameter is occurrence of disease or a particular disease in an immediate family member, thereby correlating the polymorphism with the disease.

       74.   The method of claim 71, wherein the pooled biopolymers are pooled nucleic acid molecules.

       75.   The method of claim 74, wherein the polymorphism is detected by primer oligo base extension (PROBE).

       76.   The method of 75, wherein primer oligo base extension, comprises:

       a) optionally immobilizing the nucleic acid molecules onto a solid support, to produce immobilized nucleic acid molecules;

       b) hybridizing the nucleic acid molecules with a primer oligonucleotide that is complementary to the nucleic acid molecule at a site adjacent to the polymorphism;

       c) contacting the product of step c) with composition comprising a dideoxynucleoside triphosphate or a 3'-deoxynucleoside triphosphates and a polymerase, so that only a dideoxynucleoside or 3'-deoxynucleoside triphosphate that is complementary to the polymorphism is extended onto the primer; and

       d) detecting the extended primer, thereby detecting the polymorphism in nucleic acid molecules in the pooled nucleic acids.

       77.   The method of claim 76, wherein detecting is effected by mass spectrometry.

       78.   The method of claim 71, wherein the frequency is percentage of nucleic acid molecules in the pooled nucleic acids that contain the polymorphism.

       79.   The method of claim 78, wherein the ratio is determined by obtaining mass spectra of the pooled nucleic acids.

       80.   The method of claim 72, wherein the parameter is age, thereby correlating the polymorphism with suceptibility to morbidity, early mortality or morbidity and early mortality.

81.   A method for haplotyping polymorphisms in a nucleic acid, comprising:

(a)   sorting the database of claim 8 according to a selected parameter to identify samples that match the selected parameter;

5   (b)   isolating nucleic acid from each identified sample;

(c)   optionally pooling each isolated nucleic acid;

(d)   amplifying the amount of nucleic acid;

(e)   forming single-stranded nucleic acid and splitting each single-strand into a separate reaction vessel;

10   (f)   contacting each single-stranded nucleic acid with an adaptor nucleic acid to form an adaptor complex;

(g)   contacting the adaptor complex with a nuclease and a ligase;

(h)   contacting the products of step (g) with a mixture that is capable of amplifying a ligated adaptor to produce an extended product;

15   (i)   obtaining a mass spectrum of each nucleic acid resulting from step (h) and detecting a polymorphism by identifying a signal corresponding to the extended product;

(j)   repeating steps (f) through (i) utilizing an adaptor nucleic acid able to hybridize with another adapter nucleic acid that hybridizes to a different

20   sequence on the same strand; whereby

the polymorphisms are haplotyped by detecting more than one extended product.

82.   The method of claim 1, wherein the nuclease is Fen-1.

83.   A method for haplotyping polymorphisms in a population,

25   comprising:

sorting the database of claim 8 according to a selected parameter to identify samples that match the selected parameter;

isolating a nucleic acid from each identified sample;

pooling each isolated nucleic acid;

30   optionally amplifying the amount of nucleic acid;

contacting the nucleic acid with at least one enzyme to produce fragments thereof;

-130-

obtaining a mass spectrum of the resulting fragments; whereby:

the polymorphisms are detected by detecting signals corresponding to the polymorphisms; and

the polymorphisms are haplotyped by determining from the mass

5   spectrum that the polymorphisms are located on the same strand of the nucleic acid.

84.   The method of claim 83, wherein the enzyme is a nickase.

85.   The method of claim 84, wherein the nickase is selected from the group consisting of NY2A and NYS1.

10   86.   A method for detecting methylated nucleotides within a nucleic acid sample, comprising:

splitting a nucleic acid sample into separate reaction vessels;

contacting nucleic acid in one reaction vessel with bisulfite;

amplifying the nucleic acid in each reaction vessel;

15   cleaving the nucleic acids in each reaction vessel to produce fragments thereof;

obtaining a mass spectrum of the resulting fragments from one reaction vessel and another mass spectrum of the resulting fragments from another reaction vessel; whereby:

20   cytosine methylation is detected by identifying a difference in signals between the mass spectra.

87.   The method of claim 86, wherein:

the step of amplifying is carried out in the presence of uracil; and

the step of cleaving is effected by a uracil glycosylase.

25   88.   A method for identifying a biological sample, comprising:

generating a data set indicative of the composition of the biological sample;

denoising the data set to generate denoised data;

deleting the baseline from the denoised data to generate an intermediate

30   data set;

defining putative peaks for the biological sample;

using the putative peaks to generate a residual baseline;

-131-

removing the residual baseline from the intermediate data set to generate a corrected data set;

locating, responsive to removing the residual baseline, a probable peak in the corrected data set; and

5   identifying, using the located probable peak, the biological sample;

wherein the generated biological sample data set comprises data from sense strands and antisense strands of assay fragments.

89.   The method according to claim 88, wherein identifying includes combining data from the sense strands and the antisense strands, and

10   comparing the data against expected sense strand and antisense strand values, to identify the biological

sample.

90.   The method according to claim 88, wherein identifying includes deriving a peak probability for the probable peak, in accordance with whether the

15   probable peak is from sense strand data or from antisense strand data.

91.   The method according to claim 88, wherein identifying includes deriving a peak probability for the probable peak and applying an allelic penalty in response to a ratio between a calculated area under the probable peak and a calculated expected average area under all peaks in the data set.

20   92.   A method for identifying a biological sample, comprising:

generating a data set indicative of the composition of the biological sample;

denoising the data set to generate denoised data;

deleting the baseline from the denoised data to generate an intermediate

25   data set;

defining putative peaks for the biological sample;   using the putative peaks to generate a residual baseline;

removing the residual baseline from the intermediate data set to generate a corrected data set;

30   locating, responsive to removing the residual baseline, a probable peak in the corrected data set; and

identifying, using the located probable peak, the biological sample;

wherein identifying includes deriving a peak probability for the probable peak and

applying an allelic penalty in response to a ratio between a calculated
5   area under the probable peak and a calculated expected average area under all peaks in the data set.

93.   The method according to claim 92, wherein identifying includes comparing data from probable peaks that did not receive an applied allelic penalty to determine their mass in accordance with oligonucleotide biological
10   data.

94.   The method according to claim 92, wherein the allelic penalty is not applied to probable peaks whose ratio of area under the peak to the expected area value is greater than 30%.

95.   A method for detecting a polymorphism in a nucleic acid,
15   comprising:

amplifying a region of the nucleic acid to produce an amplicon, wherein the resulting amplicon comprises one or more enzyme restriction sites;

contacting the amplicon with a restriction enzyme to produce fragments;

obtaining a mass spectrum of the resulting fragments and analyzing
20   signals in the mass spectrum by the method of claim 88; whereby:

the polymorphism is detected from the pattern of the signals.

96.   A subcollection of samples from a target population, comprising:

a plurality of samples, wherein the samples are selected from the group consisting of nucleic acids, fetal tissue, protein samples; and
25   a symbology on the containers containing the samples, wherein the symbology is representative of the source and/or history of each sample, wherein:

the target population is a healthy population that has not been selected for any disease state;
30   the collection comprises samples from the healthy population; and

the subcollection is obtained by sorting the collection according to specified parameters.

97.   The combination of claim 26, wherein the samples are selected selected from the group consisting of nucleic acids, fetal tissue, protein, tissue, body fluid, cell, seed, microbe, pathogen and reproductive tissue samples.

98.   A combination, comprising the database of claim 8 and a mass spectrometer.

99.   The combination of claim 98 that is an automated process line for analyzing biological samples.

100.   A system for high throughput processing of biological samples, comprising:

a database of claim 8, wherein the samples tested by the automated process line comprise samples from subjects in the database; and

a mass spectrometry for analysis of biopolymers in the samples.

ABSTRACT

Process and methods for creating a database of genomic samples from healthy human donors, methods that use the database to identify and correlate polymorphic genetic markers and other markers with diseases and conditions are provided.

Sample Bank

SEQUENOM
INDUSTRIAL GENOMICS

DNA Bank

Number of Samples

Caucasians

Number of Samples

Hispanics

Number of Samples

FIGURE 1

Age-Related Frequency of the 291S Allele (LPL)
Caucasians
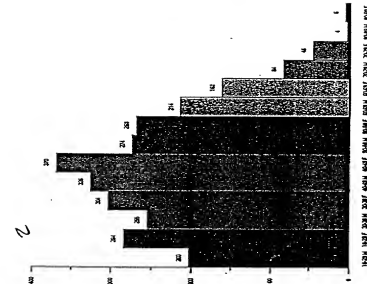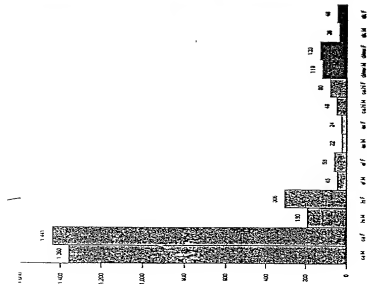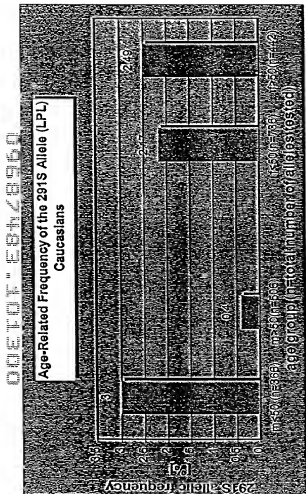
ege- and sex-distribution of the 291S allele of the lipoprotein lipase gene. A total of

436 males and 589 females were investigated.

FIGURE 2A

Age-related distribution of the 291S allele of the lipoprotein lipase gene within the

Age-Related Frequency of the 291S Allele (LPL)
Caucasian Males
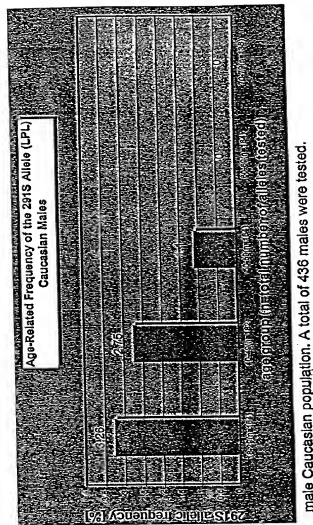
male Caucasian population. A total of 436 males were tested.
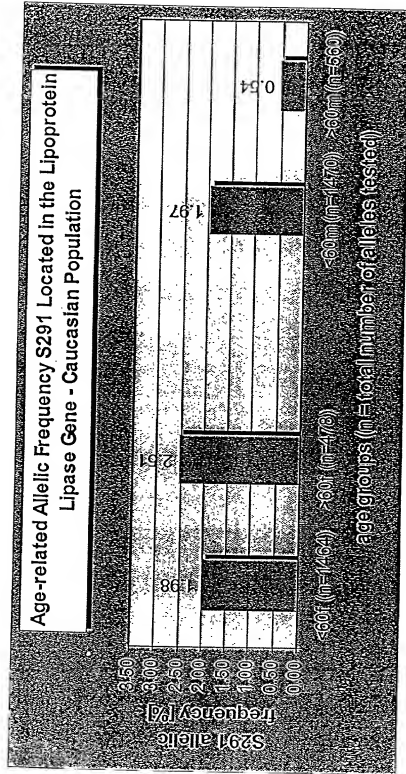
FIGURE 2B

Significance: Allele frequency of Ser291drops
from 1.97% to 0.54% in males; p=0.009

FIGURE 2C

Intelligent Genomics

SEQUENOM
INDUSTRIAL GENOMICS

Questionnaire for
Population-Based
Sample Banking

FIGURE 3

Intelligent Genomics

SEQUENOM
INDUSTRIAL GENOMICS

Sample Banks
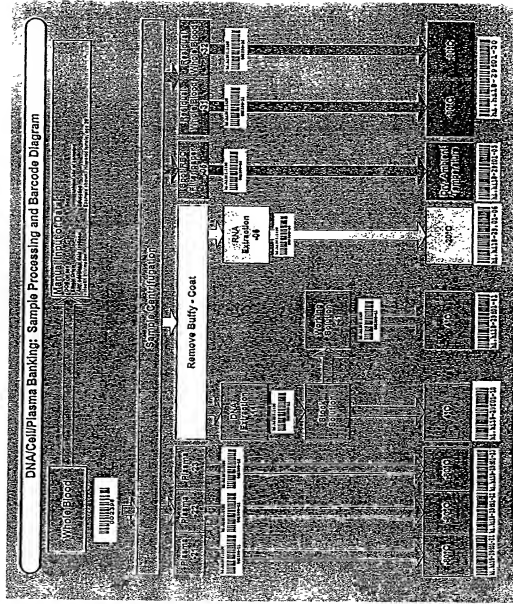
FIGURE 4

Intelligent Genomics

SEQUENOM
INDUSTRIAL GENOMICS

Age-Related Distribution of "Sick" and "Healthy" Alleles in a "Healthy" Population
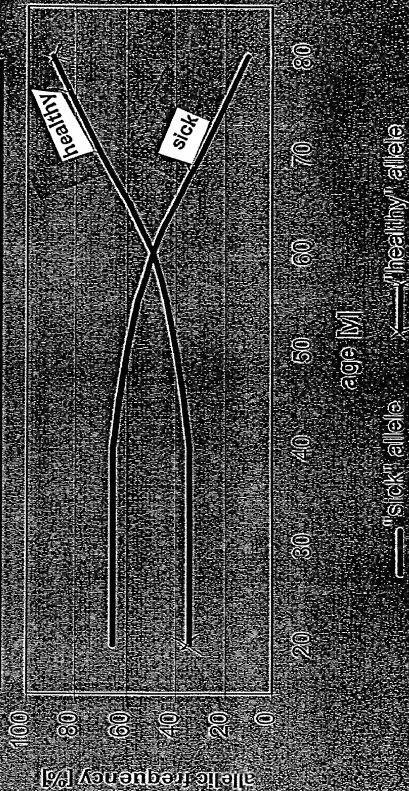
FIGURE 5

Intelligent Genomics

SEQUENOM
INDUSTRIAL GENOMICS

0037101" E34478960

Age-Dependent Distribution of ApoE Genotypes / Alleles

Centenarians   Controls (20-70 years old)

FIGURE 6

FIGURE 7A

Age-Related p53 Codon 72 Allele and Genotype Distribution
Caucasian Population

Age-related p53 Codon 72 Allele and Genotype Distribution
Caucasian Population
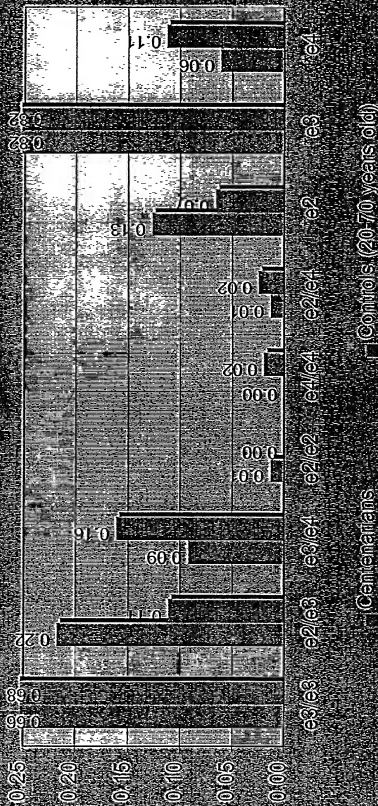
SEQUENOM
INDUSTRIAL GENOMICS
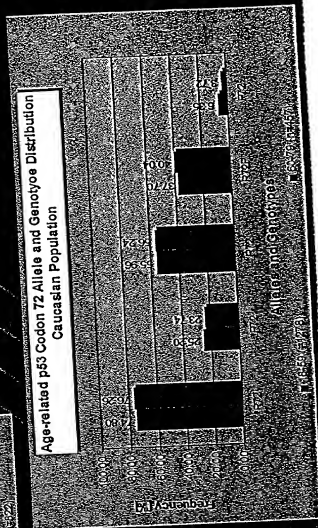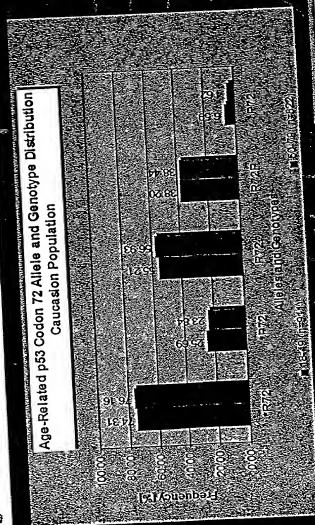
igenetics

Age-Related p53 Codon 72 Allele Distribution
Caucasian Population
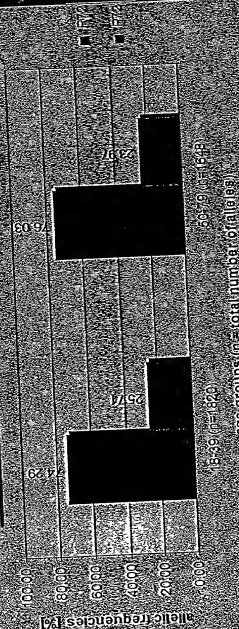
Allele Distribution among Different Ethnic Groups
Sjalander et al. Hum Hered 46: 41-48, 1996

FIGURE 7B

FIGURE 7C

i genetics

SEQUENOM
INDUSTRIAL GENOMICS

P53 PP vs. PR/RR Genotype Distribution
By Age cut point = 59

| Age Group | N | Genotype Freq (%) | |
		PP	PR/RR
18-59	1278	6.7	93.4
60-79	457	3.7	96.3

Sample Size : 1735

FIGURE 7D

Genomic Organization of the p53 Gene
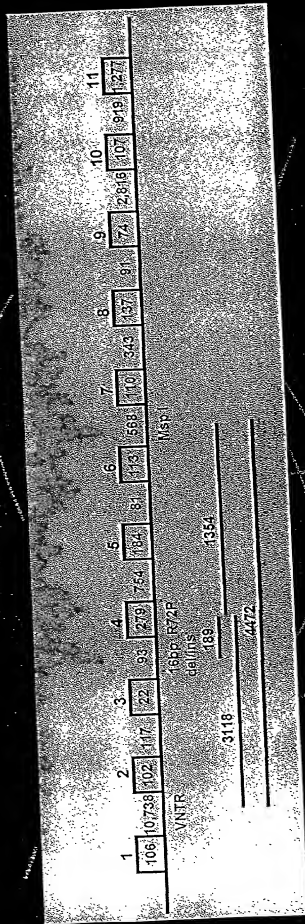
Age-Related p21 S31R Allele and Genotype Frequencies Caucasian Population
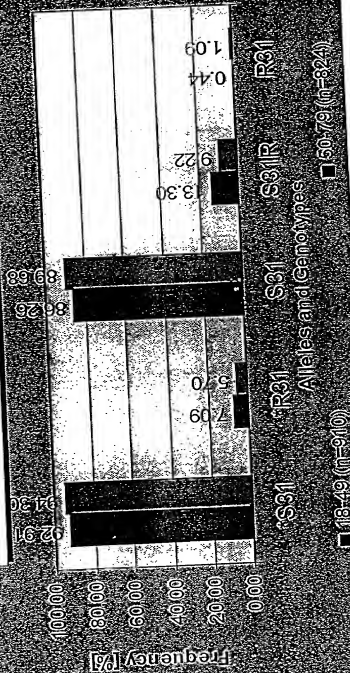
Frequency [%]

Significance: Genotype frequency of SR heterozygous drops from 13.3% to 9.2%; p=0.009

FIGURE 8

DNA MassArray™

FVII R353Q

SEQUENOM
INDUSTRIAL GENOMICS

Allelic Frequency In Pooled vs Individual Samples
FVII Allele 353Q
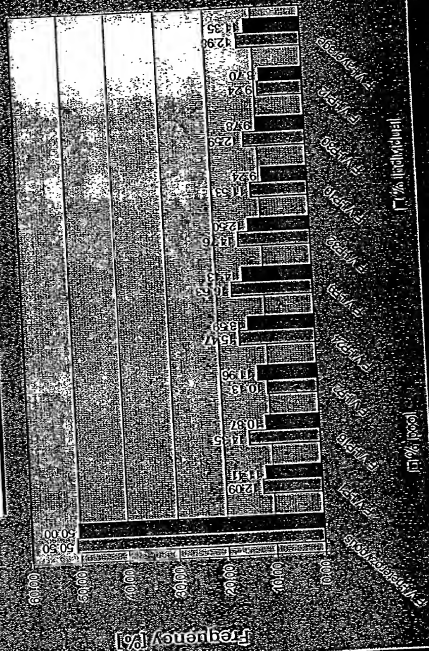
FIGURE 9

DNA MassArray™

CETP I405V
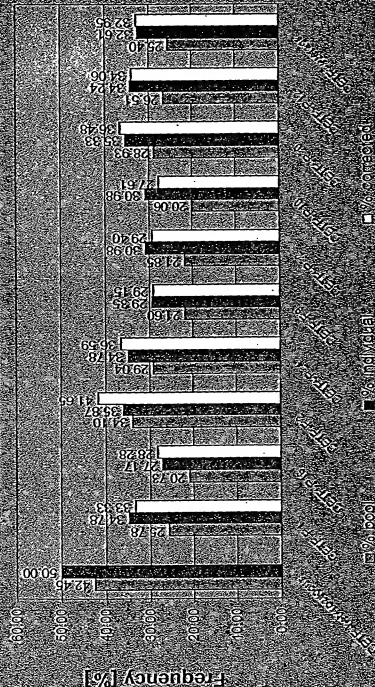
FIGURE 10

DNA MassArray™

PAI-1 4G/5G

SEQUENOM
INDUSTRIAL GENOMICS

Allelic Frequency in Pooled vs Individual Samples
Allele PAI-1 5G
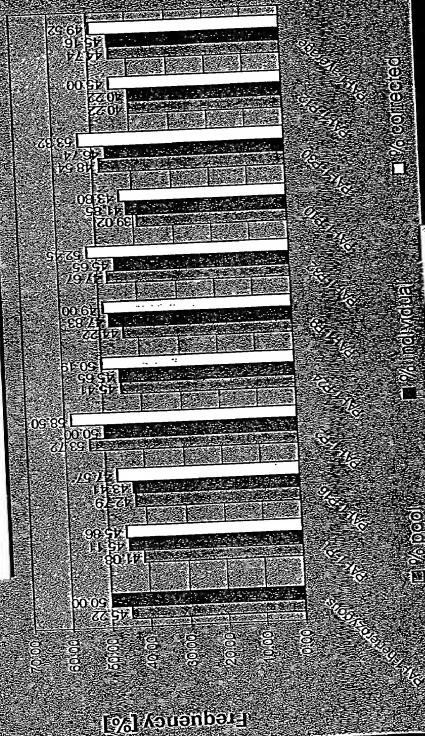
FIGURE 11

DNA MassArray™
Ethnic Diversity (PAI-1)

SEQUENOM
INDUSTRIAL GENOMICS

FIGURE 12

African   4G 28.2%   5G 71.2%

Asian   4G 65.8%   5G 34.2%

Caucasian   4G 53.5%   5G 46.5%

Hispanic   4G 26.7%   5G 73.3%
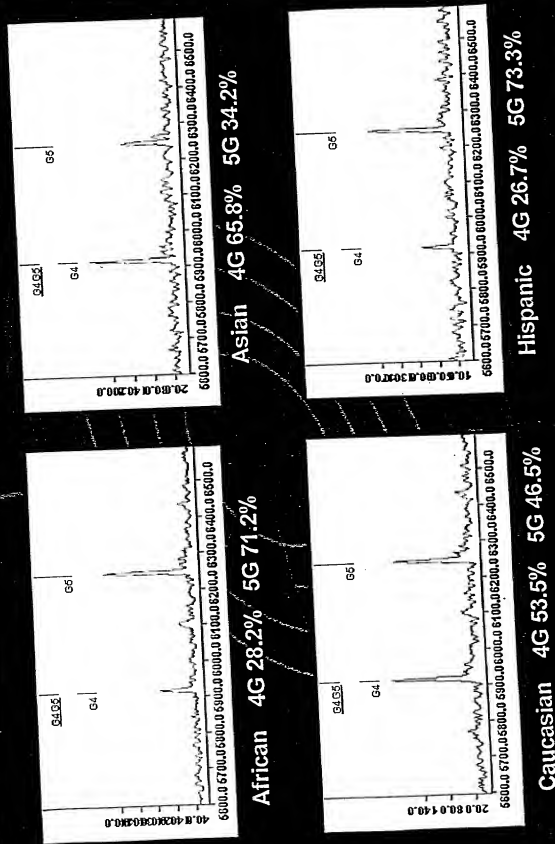
DNA MassArray™
Ethnic Diversity (CETP 405)

SEQUENOM
INDUSTRIAL GENOMICS
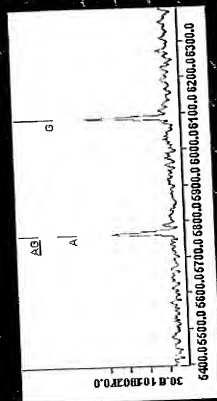
African   Ile 43.6%   Val 56.4%
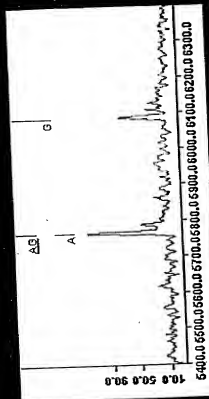
Asian   Ile 60.7%   Val 39.3%
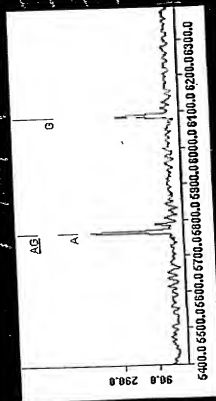
Caucasian   Ile 63.0%   Val 37.0%
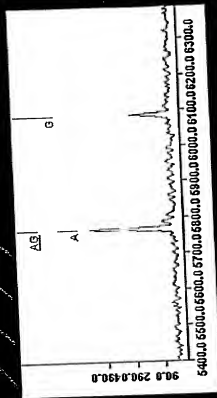
Hispanic   Ile 70.1%   Val 29.9%

FIGURE 13

DNA MassArray™
Ethnic Diversity (Factor VII 353)
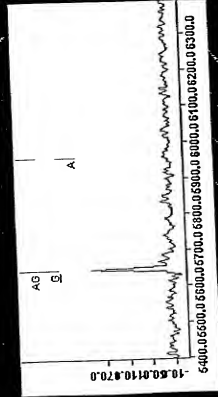
African   Arg >95% Glu <5%
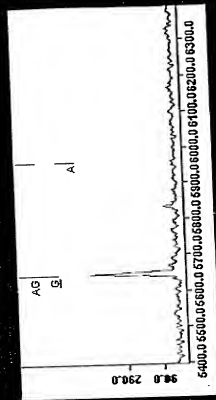
Asian   Arg >95% Glu <5%
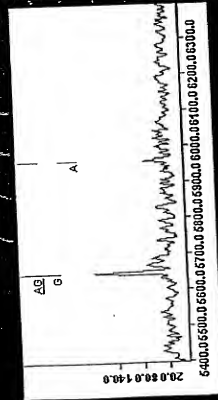
Caucasian   Arg 88.5% Glu 11.5%
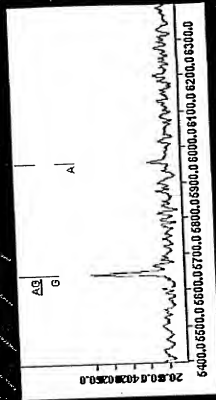
Hispanic   Arg 89.2% Glu 10.8%

FIGURE 14

DNA MassArray™
Ethnic Diversity

SEQUENOM
INDUSTRIAL GENOMICS

Ethnic diversity using pooled DNA samples

FIGURE 15

P53-Rb Pathway

Levin A. Cell, 88: 323-331, 1997

FIGURE 16

1700

FIGURE 17

START

IDENTIFY HEALTHY
MEMBERS OF A POPULATION. — 1802

OBTAIN IDENTIFYING AND HISTORICAL
INFORMATION AND DATA RELATING TO — 1804
THE IDENTIFIED MEMBERS.

ENTER THE OBTAINED
MEMBER INFORMATION AND — 1806
DATA INTO A DATABASE.

ASSOCIATE THE MEMBER
INFORMATION AND DATA — 1808
WITH AN INDEX VALUE.

CONTINUE
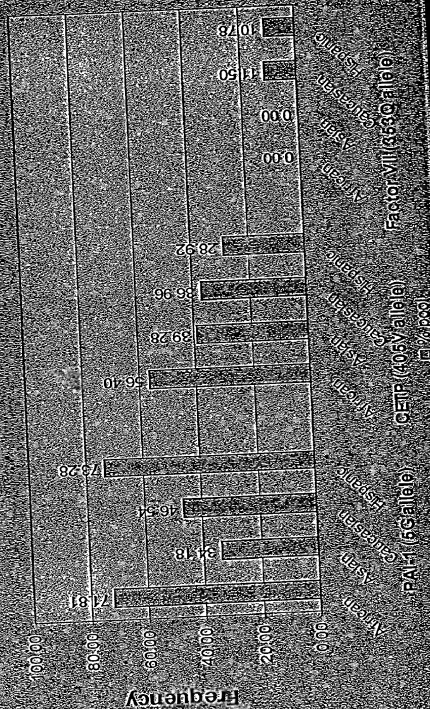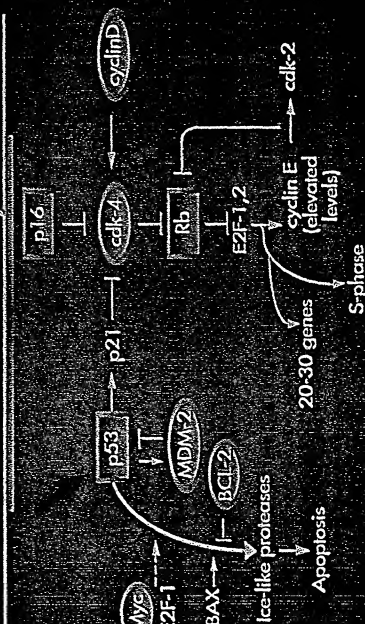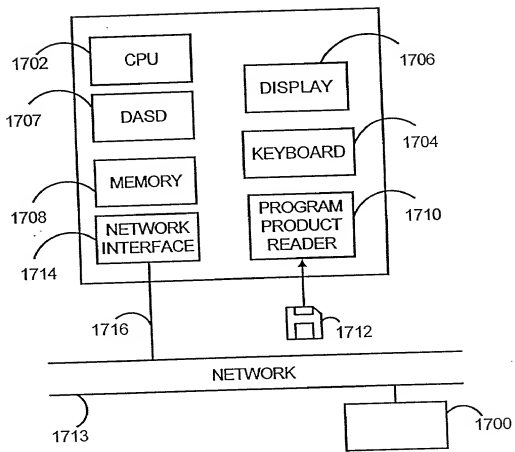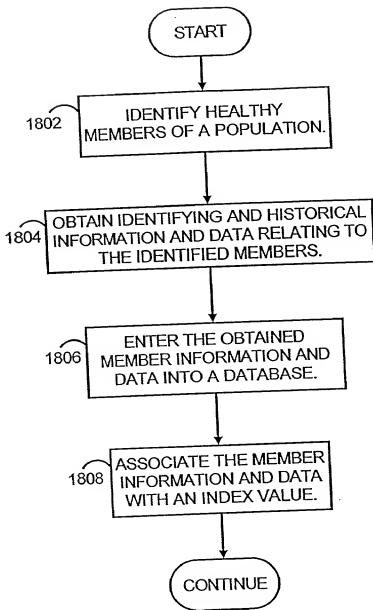
FIGURE 18

clone chromosome 17 (#48319)

Figure 19

AKAP10 Ile646Val

Figure 20

Figure 21

methionine sulfoxid reductase A (#63306)

Consent Form Signed
☐ Yes ☐ No

Time of Sample
Collection
(nearest hour, in 24 hour format)

Initials

Initials of Data Collector

Bar Code

Date of Collection

Month	Day	Year
	2 0	0 0

JAN
FEB
MAR
APR
MAY
JUN
JUL
AUG
SEP
OCT
NOV
DEC

DO NOT COMPLETE
(for data entry only)

Sample | Volume (ml)

Donor Information

Date of Birth

Month	Year
	1 9

JAN
FEB
MAR
APR
MAY
JUN
JUL
AUG
SEP
OCT
NOV
DEC

Sex:
☐ Male
☐ Female

Height
Ft. Inches

Weight (lb)

What physical activity do you do on a regular basis?
☐ Running
☐ Swimming
☐ Biking
☐ Gymnastics
☐ Other
☐ None

Are you a vegetarian?
☐ Yes
☐ No

If Female:

How many times have you been pregnant?

How many times did you give birth?

To the best of your knowledge, what is the Ethnic Origin of your:

Father   Mother
☐ ☐   Caucasian (please mark specific geographic area below if known)
☐ ☐   Northern Europe (Austria, Denmark, Finland, France, Germany, Netherlands, Norway, Sweden, Switzerland, UK)
☐ ☐   Southern Europe (Greece, Italy, Spain, Turkey)
☐ ☐   Eastern Europe (Czechoslovakia, Hungary, Poland, Russia, Yugoslavia)
☐ ☐   Middle Eastern (Israel, Egypt, Iran, Iraq, Jordan, Syria, Other Arab States)

☐ ☐   African-American

☐ ☐   Hispanic (please mark specific geographic area below if known)
☐ ☐   Mexico
☐ ☐   Central America, South America
☐ ☐   Cuba, Puerto Rico, other Caribbean

☐ ☐   Asian (please mark specific geographic area below if known)
☐ ☐   Japanese
☐ ☐   Chinese
☐ ☐   Korean
☐ ☐   Vietnamese
☐ ☐   Filipino

☐ ☐   Native American

☐ ☐   Other
☐ ☐   Don't know

In which state do you live?

A A
B B
C C
D D
E E
F F
G G
H H
I I
J J
K K
L L
M M
N N
O O
P P
Q Q
R R
S S
T T
U U
V V
W W
X X
Y Y
Z Z

How long have you lived there?
Years
0 0
1 1
2 2
3 3
4 4
5 5
6 6
7 7
8 8
9 9

What is your highest grade you completed in school?
☐ less than 8th grade
☐ 8th, 9th, 10th, or 11th grade
☐ high school graduate or equivalency
☐ some college, 2 yr degree
☐ college graduate, 4 yr degree
☐ post graduate education or degree

Mother Deceased?
☐ Yes
☐ No

If Yes at what age?

Cause of Death Mother:
☐ ≤ 29 ☐ Heart Disease
☐ 30-39 ☐ Cancer
☐ 40-49 ☐ Stroke
☐ 50-59 ☐ Accident
☐ 60-69 ☐ Suicide
☐ 70-79 ☐ Other, _____
☐ 80-89
☐ ≥ 90

Father Deceased?
☐ Yes
☐ No

If Yes at what age?

Cause of Death Father
☐ ≤ 29 ☐ Heart Disease
☐ 30-39 ☐ Cancer
☐ 40-49 ☐ Stroke
☐ 50-59 ☐ Accident
☐ 60-69 ☐ Suicide
☐ 70-79 ☐ Other, _____
☐ 80-89
☐ ≥ 90

FIGURE 22A

ave you ever smoked? ☐ Yes ☐ No

yes, for how long? ☐ Years

[checkbox columns 0-9]

Have you been hospitalized in the past 5 years for more than 6 days at a time? ☐ Yes ☐ No

If yes, how many times?
[checkboxes 1 2 3 4 5 6 7 8 9+]

For each hospitalization (if not the same) how long did you stay and for what reason?

1) Weeks: [checkboxes 1 2 3 4 5 6+]
☐ Acute disorder, including infection and thrombosis
☐ Chronic Disorder
☐ Accident
☐ Other:_____

2) Weeks: [checkboxes 1 2 3 4 5 6+]
☐ Acute disorder, including infection and thrombosis
☐ Chronic Disorder
☐ Accident
☐ Other:_____

3) Weeks: [checkboxes 1 2 3 4 5 6+]
☐ Acute disorder, including infection and thrombosis
☐ Chronic Disorder
☐ Accident
☐ Other:_____

ave you or has anyone in your immediate family (parents, brothers, sisters, or your children) had the following?
ark all that apply!

Disease	You	Mother	Father	Sister	Brother	Child
Heart Disease, including arteriosclerosis	☐	☐	☐	☐	☐	☐
Stroke	☐	☐	☐	☐	☐	☐
Hypertension	☐	☐	☐	☐	☐	☐
Abdominal blood clots	☐	☐	☐	☐	☐	☐
Diabetes, insulin dependent	☐	☐	☐	☐	☐	☐
Diabetes, not insulin-dependent (diet controlled)	☐	☐	☐	☐	☐	☐
Cancer:						
Lung & Bronchus	☐	☐	☐	☐	☐	☐
Breasts	☐	☐	☐	☐	☐	☐
Prostate	☐	☐	☐	☐	☐	☐
Colon & Rectum	☐	☐	☐	☐	☐	☐
Skin	☐	☐	☐	☐	☐	☐
Lymphoma & Leukemia	☐	☐	☐	☐	☐	☐
Other, please specify below:						
Alzheimer's Disease	☐	☐	☐	☐	☐	☐
Epilepsy	☐	☐	☐	☐	☐	☐
Schizophrenia						
Bipolar disorder (manic depression)	☐	☐	☐	☐	☐	☐
Major depression						
Chronic Inflammatory or Autoimmune Disease including Multiple Sclerosis and Rheumatoid Arthritis	☐	☐	☐	☐	☐	☐
Emphysema	☐	☐	☐	☐	☐	☐
Asthma						
Other, please specify below:						

Do you take prescription drugs on a regular basis?   ☐ Yes ☐ No

f yes, please specify below:

ave you ever donated blood before?   ☐ Yes ☐ No

f yes, how many times: Number of Times
[number columns]

Additional health information details you would like to provide:
_____ ___ ___
_____
_____
_____

Do you drink any kind of alcoholic beverage?
Never   Hardly ever
Less than 3 times per week   3 or more times per week
Daily

FIGURE 22B

Collection Information

Consent Form Signed
☐ Yes ☐ No

Date of Collection

Month	Day	Year
JAN		2 0 0
FEB		
MAR		
APR		
MAY		
JUN		
JUL		
AUG		
SEP		
OCT		
NOV		
DEC		

Time of Sample Collection
(nearest hour, in 24 hour clock format)

Initials

Initials of Data Collector

(DO NOT COMPLETE; for data entry only)
Sample:
☐ Received
☐ Lost
☐ Broken

Volume (ml)

Bar Code

Donor Information

Date of Birth

Month	Year
JAN	1 9
FEB	
MAR	
APR	
MAY	
JUN	
JUL	
AUG	
SEP	
OCT	
NOV	
DEC	

Sex:
☐ Male
☐ Female

Height (Ft. Inches)

Weight (lb)

What physical activity do you do on a regular basis?
☐ Running
☐ Swimming
☐ Biking
☐ Gymnastics
☐ Other
☐ None

Are you a vegetarian?
☐ Yes
☐ No

If Female:
How many times have you been pregnant?
How many times did you give birth?

In which state do you live?

To the best of your knowledge, what is the Ethnic Origin of your:

Father Mother
☐ ☐ Caucasian (please mark specific geographic area below if known)
☐ ☐ Northern Europe (Austria, Denmark, Finland, France, Germany, Netherlands, Norway, Sweden, Switzerland, UK)
☐ ☐ Southern Europe (Greece, Italy, Spain, Turkey)
☐ ☐ Eastern Europe (Czechoslovakia, Hungary, Poland, Russia, Yugoslavia)
☐ ☐ Middle Eastern (Israel, Egypt, Iran, Iraq, Jordan, Syria, Other Arab States)
☐ ☐ Other
☐ ☐ Don't know

How long have you lived there? Years

How many years have you been smoking? Years

Did you quit smoking?
☐ Yes
☐ No

If yes, how many years ago? Years

How many cigarettes do/did you smoke per day? Cigarettes

Do you have Lung Emphysema?
☐ Yes
☐ No

If yes, for how long? Years

Continue on back →

FIGURE 22C

...l completed in school?
- less than 8th grade
- 8th, 9th, 10th, or 11th grade
- high school graduate or equivalency
- some college, 2 yr degree
- college graduate, 4 yr degree
- post graduate education or degree

☐ Yes ☐ No

If Yes at what age?
- ☐ ≤ 29
- ☐ 30-39
- ☐ 40-49
- ☐ 50-59
- ☐ 60-69
- ☐ 70-79
- ☐ 80-89
- ☐ ≥ 90

- ☐ Heart Disease
- ☐ Cancer
- ☐ Stroke
- ☐ Accident
- ☐ Suicide
- ☐ Other_____

☐ Yes ☐ No

If Yes at what age?
- ☐ ≤ 29
- ☐ 30-39
- ☐ 40-49
- ☐ 50-59
- ☐ 60-69
- ☐ 70-79
- ☐ 80-89
- ☐ ≥ 90

- ☐ Heart Disease
- ☐ Cancer
- ☐ Stroke
- ☐ Accident
- ☐ Suicide
- ☐ Other_____

Health Information

Have you or has anyone in your immediate family (parents, brothers, sisters, or your children) had the following? Mark all that apply!

Disease	You	Mother	Father	Sister	Brother	Child
Heart Disease	☐	☐	☐	☐	☐	☐
Stroke	☐	☐	☐	☐	☐	☐
Hypertension	☐	☐	☐	☐	☐	☐
Abdominal blood clots	☐	☐	☐	☐	☐	☐
Diabetes, insulin dependent	☐	☐	☐	☐	☐	☐
Diabetes, not insulin-dependent (diet controlled)	☐	☐	☐	☐	☐	☐
Cancer:						
Lung & Bronchus	☐	☐	☐	☐	☐	☐
Breasts	☐	☐	☐	☐	☐	☐
Prostate	☐	☐	☐	☐	☐	☐
Colon & Rectum	☐	☐	☐	☐	☐	☐
Skin	☐	☐	☐	☐	☐	☐
Lymphoma & Leukemia	☐	☐	☐	☐	☐	☐
Other, please specify below:						
Alzheimer's Disease	☐	☐	☐	☐	☐	☐
Epilepsy	☐	☐	☐	☐	☐	☐
Schizophrenia	☐	☐	☐	☐	☐	☐
Bipolar disorder (manic depression)	☐	☐	☐	☐	☐	☐
Major depression	☐	☐	☐	☐	☐	☐
Chronic Inflammatory or Autoimmune Disease including Multiple Sclerosis and Rheumatoid Arthritis						
Emphysema	☐	☐	☐	☐	☐	☐
Asthma	☐	☐	☐	☐	☐	☐
Other, please specify below:						

Do you take prescription drugs on a regular basis? ☐ Yes ☐ No

If yes, please specify below:

Have you ever donated blood before? ☐ Yes ☐ No

If yes, how many times: Number of Times

☐☐☐

Have you been hospitalized in the past 5 years for more than 6 days at a time?
☐ Yes ☐ No

If yes, how many times?
☐☐☐☐☐☐☐

For each hospitalization (if not the same) how long did you stay and for what reason?

1) Weeks: ☐☐☐☐☐☐
- ☐ Acute disorder, including infection and thrombosis
- ☐ Chronic Disorder
- ☐ Accident
- ☐ Other:_____

2) Weeks: ☐☐☐☐☐☐
- ☐ Acute disorder, including infection and thrombosis
- ☐ Chronic Disorder
- ☐ Accident
- ☐ Other:_____

3) Weeks: ☐☐☐☐☐☐
- ☐ Acute disorder, including infection and thrombosis
- ☐ Chronic Disorder
- ☐ Accident
- ☐ Other:_____

☐☐☐
☐☐☐
☐☐☐
☐☐☐
☐☐☐
☐☐☐
☐☐☐

Do you drink any kind of alcoholic beverage?
☐ Never   ☐ Hardly ever
☐ Less than 3 times per week   ☐ 3 or more times per week
☐ Daily

Additional health information details you would like to provide:

_____
_____
_____
_____

FOR OFFICE USE ONLY

☐ ☐
☐ ■
☐ ☐
☐ ☐
☐ ☐

FIGURE 22D

RECEIVED DATA FROM MASS SPECTROMETER — 602

FIND THE HEIGHT (H) OF EACH PEAK — 604

BEGIN DENOISE — 605

EXTRAPOLATE NOISE PROFILE — 606

FIND NOISE OF EACH PEAK — 608

CALCULATE SIGNAL TO NOISE RATIO — 610

FIND RESIDUAL ERROR — 612

PERFORM SENSE STRAND PROCESSING — 630

PERFORM ANTISENSE STRAND PROCESSING — 640

COMBINE DATA AND COMPARE TO EXPECTED DATABASE VALUES — 650

CALL GENOTYPE — 660

FIGURE 23

FIGURE 24

40 — RECEIVE DATA

35

45 — DENOISE
- Wavelet Transformation
- Noise Profile
- Apply Noise Profile to each Stage

50 — BASELINE CORRECTION
- Generate Sparse Data Set
- Shift Signal
- Generate Peak-Free Signal
- Generate Baseline
- Subtract Baseline From Signal

55 — COMPRESSION

60 — PEAK FITTING
- Mass Shift
- Find width of Strongest Peak
- Delete an Area Around Each Peak
- Generate a Residual Baseline
- Subtract Residual Baseline From Signal
- Determine Height of Each Peak

65 — GENOTYPE CALLING
- Determine Peak Probabilities
- Calling Genotypes

FIGURE 25

FIGURE 26

FIGURE 27

95

FIGURE 28

97
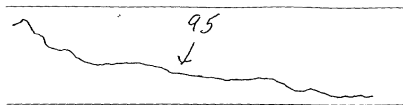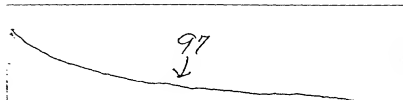
Exp fitting

$a_0 + a_1 \exp^-(a_2 m)$

FIGURE 29
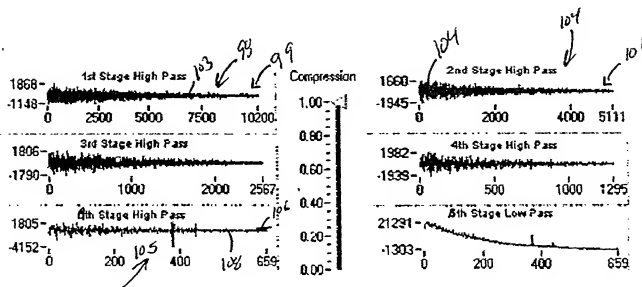
FIGURE 30

Stage 0 - Hi

Threshold 0=4XNoiseProfile

Stage 1 - Hi
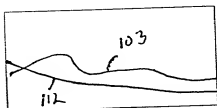
Threshold 1=2XNoiseProfile

Stage 2 - Hi
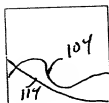
Threshold 2=1XNoiseProfile

Stage n - Hi
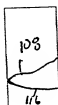
Threshold n=(1/2^{n-2})XNoiseProfile

Stage n - Lo
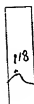
FIGURE 31
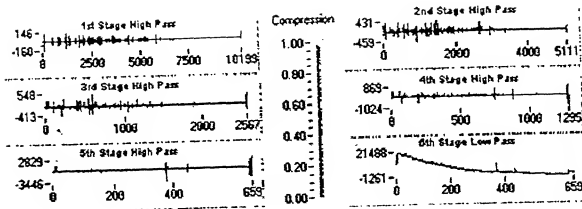
FIGURE 32

$$\text{Signal (t)} = \frac{\text{(Start 0(t) + Start 1(t) + Start 2 (t)... + Start 23 (t))}}{24}$$

FIGURE 33   SHIFT SIGNAL TO ACCOUNT FOR
                 VARIATIONS DUE TO STARTING POINT

FIGURE 34

130

FIGURE 35

150

147
145
149
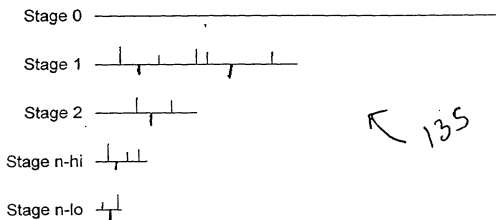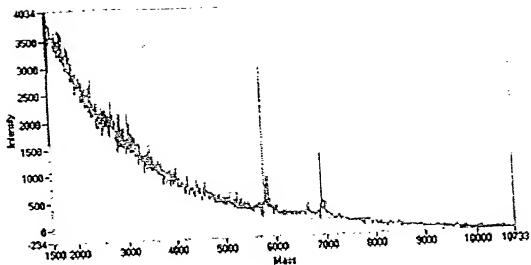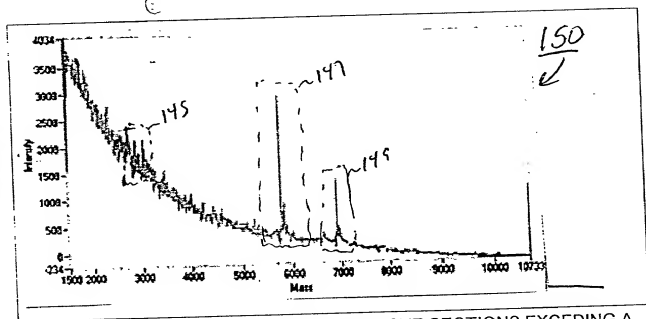
FIG. 13 - TAKE A MOVING AVERAGE, REMOVE SECTIONS EXCEEDING A
THRESHOLD

FIGURE 36
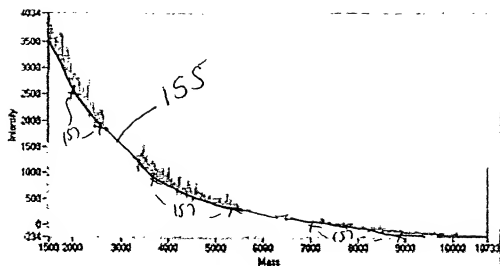
FIGURE 37

FIND MINIMA IN REMAINING SIGNALS AND CONNECT TO
FORM A PEAK FREE SIGNAL

| DO A WAVELET TRANSFORMATION ON THE PEAK FREE SIGNAL | DELETE ALL STAGES, LEAVING ON THE N-LO STAGE | RECONSTRUCT A SIGNAL FROM N-LO | GENERATE BASELINE CORRECTION SIGNAL |
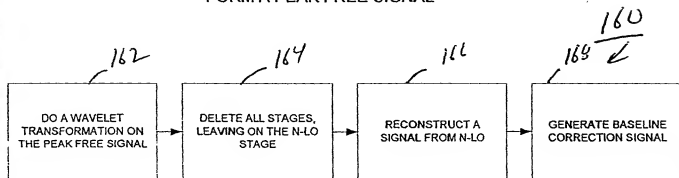
GENERATE BASELINE CORRECTION
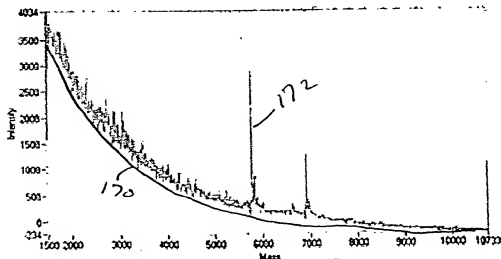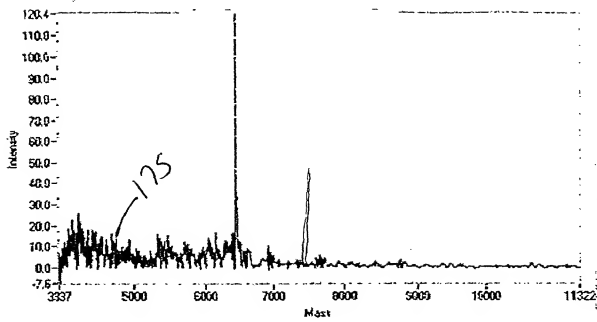
FIGURE 38

FIGURE 39

FIGURE 40

NON-0 COEFFICIENTS	VALUE	INTERMEDIATE	RELATIVE
100	25	100.025	100.025
150	220	150.220	50.220
500	.1	500.0001	350.0001
10,050	800	10,050.8	9550.8
10,075	890	10,075.89	25.89
11,125	910	11,125.91	150.91
12,100	1000 (MAX)	12,100.99999	975.99999
13,250	940	13,250.94	1150.94

FIGURE 41

201 RECEIVE DATA FILE

202 ARRANGE AN X-Y ARRAY OF COEFFICIENTS

203 DETERMINE STARTING POINT AND MAXIMUM DATA VALUE

204 GENERATE INTERMEDIATE REAL NUMBER

Whole Portion is X axis number

Decimal Portion is fractional Data Value

205 CALCULATE COMPRESSED VALUE

200
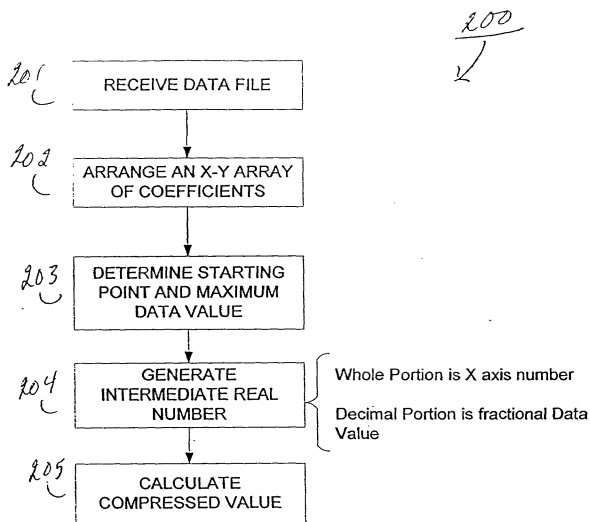
FIGURE 42

FIGURE 43

FIGURE 44

230 218 220 232 222 234 224

←247→|←500→|247 238 |²⁴⁰|←500→| 240 |²⁴⁰|←500→|245
²⁴⁶ **FIGURE 45**

247 245

FIGURE 46

250
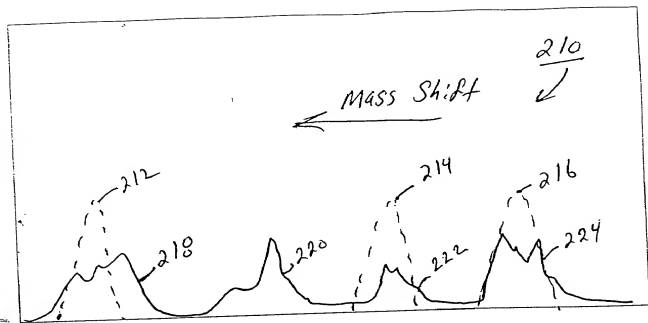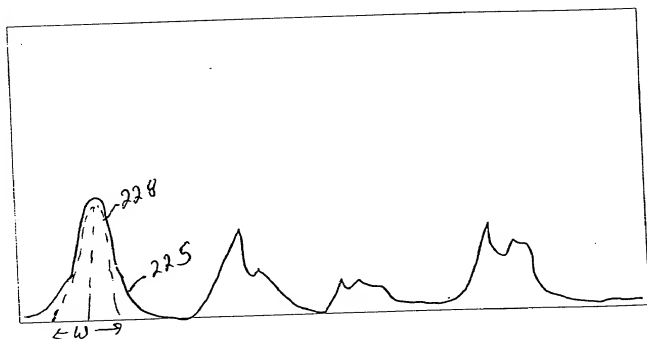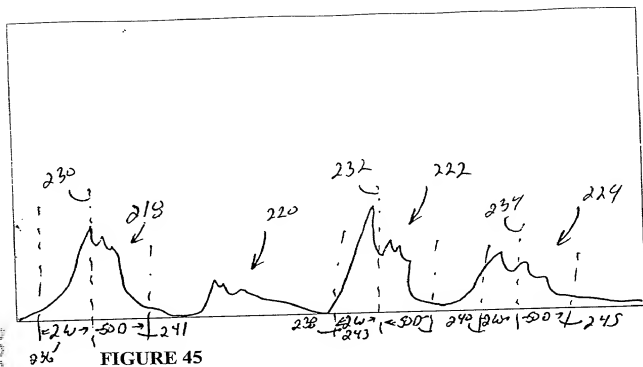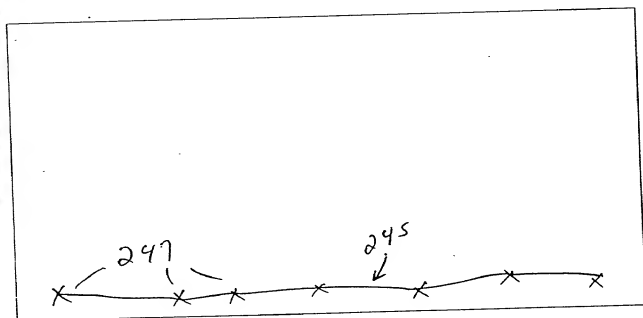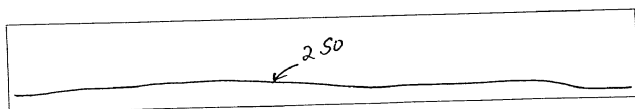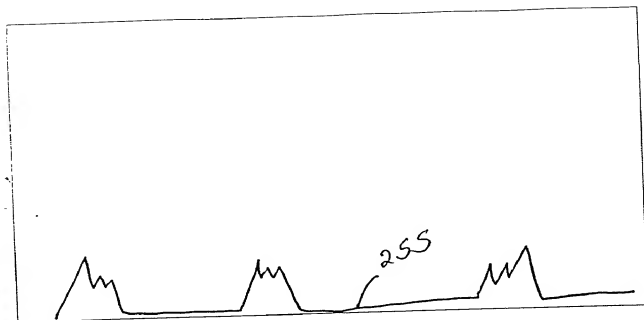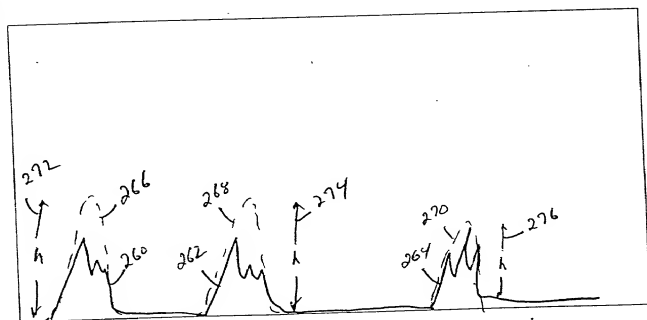
FIGURE 47

FIGURE 48

FIGURE 49
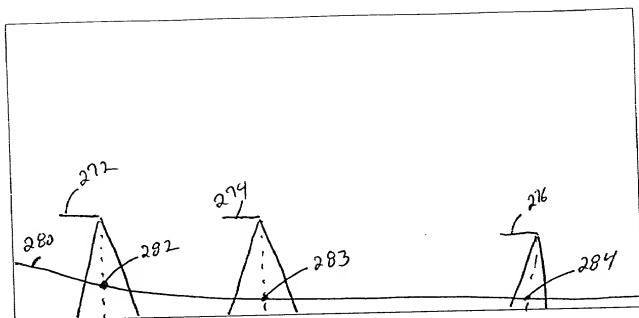
FIGURE 50
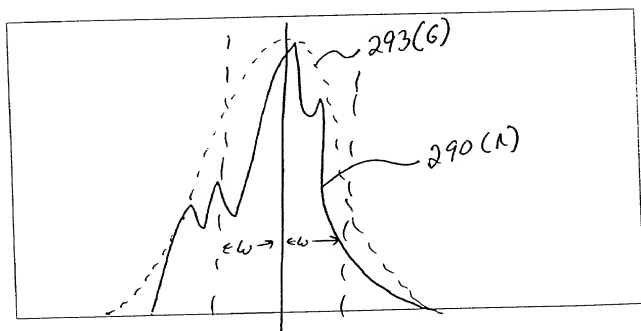
FIGURE 51
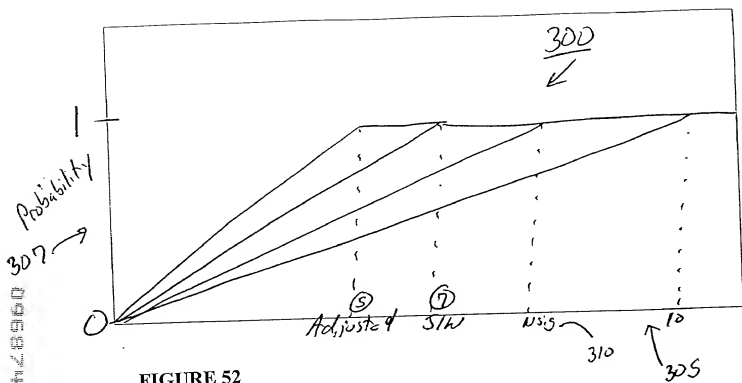
FIGURE 52
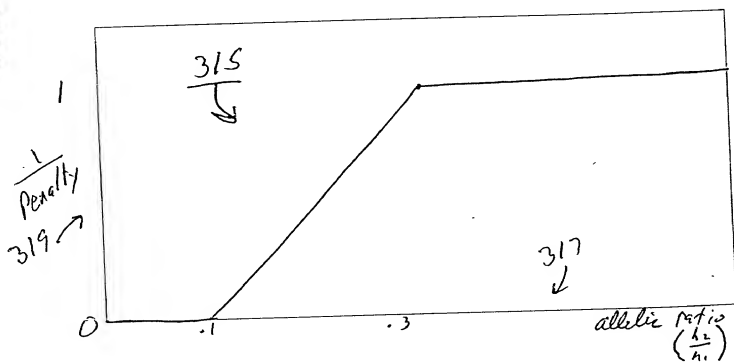
FIGURE 53
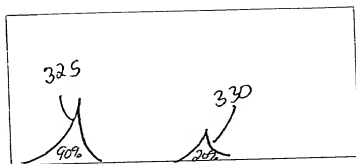
329

PROBABILITY OF GG EXISTING:
P(GG) = P(G) * P(1-C)
        =90% * (100% - 20%)
        =90% * 80%
        =72%

331

PROBABILITY OF CC EXISTING:
P(CC) = P(C) * P(1-G)
        =20% * (100% - 90%)
        =20% * 10%
        = 2%

333

PROBABILITY OF GC EXISTING:
P(GC) = P(G) * P(C)
        =90% *20%
        =18%

325

330

90%   20%

FIGURE 54

335

Call

339

No call

335

3   2.5   9   337

m   Ratio $\left(\frac{P_1}{P_2}\right)$

FIGURE 55

400

FIND H OF EACH PEAK — 402

EXTRAPOLATE NOISE PROFILE — 404

FIND NOISE OF EACH PEAK — 406

CALCULATE A SIGNAL TO NOISE RATIO — 410

FIND A RESIDUAL ERROR — 412

CALCULATE AN ADJUSTED SIGNAL TO NOISE RATIO — 414

DEVELOP A PROBABILITY PROFILE — 416

DETERMINE PEAK PROBABILITIES — 418

DETERMINE AN ALLELIC PENALTY — 420

ADJUST PEAK PROBABILITY BY ALLELIC PENALTY — 422

CALCULATE GENOTYPE PROBABILITIES — 424

TEST RATIO OF GENOTYPE PROBABILITIES — 426

CALL GENOTYPE — 428

FIGURE 56

RECEIVE DATA FROM MASS SPECTROMETER — 502

FIND THE HEIGHT (H) OF EACH PEAK — 504

BEGIN DENOISE — 505

EXTRAPOLATE NOISE PROFILE — 506

FIND NOISE OF EACH PEAK — 508

CALCULATE SIGNAL TO NOISE RATIO — 510

FIND RESIDUAL ERROR — 512

CALCULATE AN ADJUSTED SIGNAL TO NOISE RATIO — 514

DEVELOP A PROBABILITY PROFILE — 516

DETERMINE PEAK PROBABILITIES — 518

DETERMINE ALLELIC PENALTY FOR EACH PEAK — 520

IDENTIFICATION OF PEAK(S) THAT HAVE ALLELIC RATIOS BELOW THRESHOLD — 522

DETERMINE MASS OF PEAK — 524

FIGURE 57

RATIO OF AREA UNDER PEAK

FIGURE 58

<110> Braun *et al.*

<120> METHODS FOR GENERATING DATABASES AND DATABASES FOR IDENTIFYING POLYMORPHIC GENETIC MARKERS

<130> 24736-2033

<140> Not Yet Assigned
<141> 2000-10-13

<150> 60/217,658
<151> 2000-07-10

<150> 60/159,176
<151> 1999-10-13

<150> 60/217,251
<151> 2000-07-10

<150> 09/663,968
<151> 2000-09-19

<160> 118

<170> FastSEQ for Windows Version 4.0

<210> 1
<211> 361
<212> DNA
<213> Homo Sapien

<400> 1
```
ctgaggacct ggtcctctga ctgctctttt cacccatcta cagtcccct tgccgtccca      60
agcaatggat gatttgatgc tgtccccgga cgatattgaa caatggttca ctgaagaccc     120
aggtccagat gaagctccca gaatgccaga ggctgctccc cgcgtggccc ctgcaccagc     180
agctcctaca ccggcggccc ctgcaccagc ccctcctgg ccctgtcat cttctgtccc      240
ttcccagaaa acctaccagg gcagctacgg tttccgtctg ggcttcttgc attctgggac     300
agccaagtct gtgacttgca cggtcagttg ccctgagggg ctggcttcca tgagacttca     360
a                                                                      361
```

<210> 2
<211> 44
<212> DNA
<213> Artificial Sequence

<220>
<223> Oligonucleotide Primer

<400> 2
```
cccagtcacg acgttgtaaa acgctgagga cctggtcctc tgac                       44
```

<210> 3
<211> 42
<212> DNA
<213> Artificial Sequence

<220>
<223> Oligonucleotide Primer

<400> 3
```
agcggataac aatttcacac aggttgaagt ctcatggaag cc                         42
```

<210> 4
<211> 17

1

```
<212> DNA
<213> Artificial Sequence

<220>
<223> Probe

<400> 4
gccagaggct gctcccc          .                                    17

<210> 5
<211> 17
<212> DNA
<213> Artificial Sequence

<220>
<223> Probe

<400> 5
gccagaggct gctcccc                                               17

<210> 6
<211> 19
<212> DNA
<213> Artificial Sequence

<220>
<223> Probe

<400> 6
gccagaggct gctccccgc                                             19

<210> 7
<211> 18
<212> DNA
<213> Artificial Sequence

<220>
<223> Probe

<400> 7
gccagaggct gctccccc                                              18

<210> 8
<211> 161
<212> DNA
<213> Homo Sapien

<400> 8
gtccgtcaga acccatgcgg cagcaaggcc tgccgccgcc tcttcggccc agtggacagc    60
gagcagctga gccgcgactg tgatgcgcta atggcgggct gcatccagga ggcccgtgag   120
cgatggaact tcgactttgt caccgagaca ccactggagg g                       161

<210> 9
<211> 43
<212> DNA
<213> Artificial Sequence

<220>
<223> Oligonucleotide Primer

<400> 9
cccagtcacg acgttgtaaa acggtccgtc agaacccatg cgg                     43

<210> 10
<211> 44
<212> DNA
<213> Artificial Sequence
```

2

<220>
<223> Oligonucleotide Primer

<400> 10
agcggataac aatttcacac aggctccagt ggtgtctcgg tgac   44

<210> 11
<211> 15
<212> DNA
<213> Artificial Sequence

<220>
<223> Oligonucleotide Primer

<400> 11
cagcgagcag ctgag   15

<210> 12
<211> 15
<212> DNA
<213> Artificial Sequence

<220>
<223> Probe

<400> 12
cagcgagcag ctgag   15

<210> 13
<211> 16
<212> DNA
<213> Artificial Sequence

<220>
<223> Probe

<400> 13
cagcgagcag ctgagc   16

<210> 14
<211> 17
<212> DNA
<213> Artificial Sequence

<220>
<223> Probe

<400> 14
cagcgagcag ctgagac   17

<210> 15
<211> 205
<212> DNA
<213> Homo Sapien

<400> 15
gcgctccatt catctcttca tcgactctct gttgaatgaa gaaaatccaa gtaaggccta   60
caggtgcagt tccaaggaag cctttgagaa agggctctgc ttgagttgta gaaagaaccg   120
ctgcaacaat ctgggctatg agatcaataa agtcagagcc aaaagaagca gcaaaatgta   180
cctgaagact cgttctcaga tgccc   205

<210> 16
<211> 42
<212> DNA
<213> Artificial Sequence

<220>

4

<400> 22
gtgccggcta ctcggatggc agcaaggact cctgcaaggg ggacagtgga ggcccacatg   60

<210> 23
<211> 60
<212> DNA
<213> Homo sapien

<400> 23
ccacccacta ccggggcacg tggtacctga cgggcatcgt cagctggggc cagggctgcg   60

<210> 24
<211> 42
<212> DNA
<213> Artificial Sequence

<220>
<223> Oligonucleotide primer

<400> 24
cccagtcacg acgttgtaaa acgatggcag caaggactcc tg   42

<210> 25
<211> 18
<212> DNA
<213> Artificial Sequence

<220>
<223> Oligonucleotide primer

<400> 25
cacatgccac ccactacc   18

<210> 26
<211> 43
<212> DNA
<213> Artificial Sequence

<220>
<223> Oligonucleotide primer

<400> 26
agcggataac aatttcacac aggtgacgat gcccgtcagg tac   43

<210> 27
<211> 15
<212> DNA
<213> Artificial Sequence

<220>
<223> Probe

<400> 27
atgccaccca ctacc   15

<210> 28
<211> 19
<212> DNA
<213> Artificial Sequence

<220>
<223> Probe

<400> 28
cacatgccac ccactaccg   19

<210> 29

5

```
<211> 20
<212> DNA
<213> Artificial Sequence

<220>
<223> Probe

<400> 29
cacatgccac ccactaccag                                                    20


<210> 30
<211> 23
<212> DNA
<213> Artificial Sequence

<220>
<223> Probe

<400> 30
agcggataac aatttcacac agg                                                23


<210> 31
<211> 2363
<212> DNA
<213> Homo Sapien

<220>
<221> CDS
<222> (138)...(2126)
<223> AKAP-10

<300>
<308> GenBank AF037439
<309> 1997-12-21

<400> 31
gcggcttgtt gataatatgg cggctggagc tgcctgggca tcccgaggag gcggtggggc     60
ccactcccgg aagaagggtc cctttcgcg ctagtgcagc ggccctctg gacccggaag     120
tccgggccgg ttgctga atg agg gga gcc ggg ccc tcc ccg cgc cag tcc         170
                     Met Arg Gly Ala Gly Pro Ser Pro Arg Gln Ser
                      1               5                   10

ccc cgc acc ctc cgt ccc gac ccg ggc ccc gcc atg tcc ttc ttc cgg       218
Pro Arg Thr Leu Arg Pro Asp Pro Gly Pro Ala Met Ser Phe Phe Arg
                15                  20                  25

cgg aaa gtg aaa ggc aaa gaa caa gag aag acc tca gat gtg aag tcc       266
Arg Lys Val Lys Gly Lys Glu Gln Glu Lys Thr Ser Asp Val Lys Ser
        30                  35                  40

att aaa gct tca ata tcc gta cat tcc cca caa aaa agc act aaa aat       314
Ile Lys Ala Ser Ile Ser Val His Ser Pro Gln Lys Ser Thr Lys Asn
    45                  50                  55

cat gcc ttg ctg gag gct gca gga cca agt cat gtt gca atc aat gcc       362
His Ala Leu Leu Glu Ala Ala Gly Pro Ser His Val Ala Ile Asn Ala
60                  65                  70                  75

att tct gcc aac atg gac tcc ttt tca agt agc agg aca gcc aca ctt       410
Ile Ser Ala Asn Met Asp Ser Phe Ser Ser Ser Arg Thr Ala Thr Leu
                80                  85                  90

aag aag cag cca agc cac atg gag gct gct cat ttt ggt gac ctg ggc       458
Lys Lys Gln Pro Ser His Met Glu Ala Ala His Phe Gly Asp Leu Gly
            95                  100                 105

aga tct tgt ctg gac tac cag act caa gag acc aaa tca agc ctt tct       506
```

6

```
Arg Ser Cys Leu Asp Tyr Gln Thr Gln Glu Thr Lys Ser Ser Leu Ser
    110                     115                 120

aag acc ctt gaa caa gtc ttg cac gac act att gtc ctc cct tac ttc     554
Lys Thr Leu Glu Gln Val Leu His Asp Thr Ile Val Leu Pro Tyr Phe
125                     130                 135

att caa ttc atg gaa ctt cgg cga atg gag cat ttg gtg aaa ttt tgg     602
Ile Gln Phe Met Glu Leu Arg Arg Met Glu His Leu Val Lys Phe Trp
140                 145                 150                 155

tta gag gct gaa agt ttt cat tca aca act tgg tcg cga ata aga gca     650
Leu Glu Ala Glu Ser Phe His Ser Thr Thr Trp Ser Arg Ile Arg Ala
                160                 165                 170

cac agt cta aac aca atg aag cag agc tca ctg gct gag cct gtc tct     698
His Ser Leu Asn Thr Met Lys Gln Ser Ser Leu Ala Glu Pro Val Ser
            175                 180                 185

cca tct aaa aag cat gaa act aca gcg tct ttt tta act gat tct ctt     746
Pro Ser Lys Lys His Glu Thr Thr Ala Ser Phe Leu Thr Asp Ser Leu
        190                 195                 200

gat aag aga ttg gag gat tct ggc tca gca cag ttg ttt atg act cat     794
Asp Lys Arg Leu Glu Asp Ser Gly Ser Ala Gln Leu Phe Met Thr His
    205                 210                 215

tca gaa gga att gac ctg aat aat aga act aac agc act cag aat cac     842
Ser Glu Gly Ile Asp Leu Asn Asn Arg Thr Asn Ser Thr Gln Asn His
220                 225                 230                 235

ttg ctg ctt tcc cag gaa tgt gac agt gcc cat tct ctc cgt ctt gaa     890
Leu Leu Leu Ser Gln Glu Cys Asp Ser Ala His Ser Leu Arg Leu Glu
                240                 245                 250

atg gcc aga gca gga act cac caa gtt tcc atg gaa acc caa gaa tct     938
Met Ala Arg Ala Gly Thr His Gln Val Ser Met Glu Thr Gln Glu Ser
            255                 260                 265

tcc tct aca ctt aca gta gcc agt aga aat agt ccc gct tct cca cta     986
Ser Ser Thr Leu Thr Val Ala Ser Arg Asn Ser Pro Ala Ser Pro Leu
        270                 275                 280

aaa gaa ttg tca gga aaa cta atg aaa tct ata gaa caa gat gca gtg    1034
Lys Glu Leu Ser Gly Lys Leu Met Lys Ser Ile Glu Gln Asp Ala Val
    285                 290                 295

aat act ttt acc aaa tat ata tct cca gat gct gct aaa cca ata cca    1082
Asn Thr Phe Thr Lys Tyr Ile Ser Pro Asp Ala Ala Lys Pro Ile Pro
300                 305                 310                 315

att aca gaa gca atg aga aat gac atc ata gca agg att tgt gga gaa    1130
Ile Thr Glu Ala Met Arg Asn Asp Ile Ile Ala Arg Ile Cys Gly Glu
                320                 325                 330

gat gga cag gtg gat ccc aac tgt ttc gtt ttg gca cag tcc ata gtc    1178
Asp Gly Gln Val Asp Pro Asn Cys Phe Val Leu Ala Gln Ser Ile Val
            335                 340                 345

ttt agt gca atg gag caa gag cac ttt agt gag ttt ctg cga agt cac    1226
Phe Ser Ala Met Glu Gln Glu His Phe Ser Glu Phe Leu Arg Ser His
        350                 355                 360

cat ttc tgt aaa tac cag att gaa gtg ctg acc agt gga act gtt tac    1274
His Phe Cys Lys Tyr Gln Ile Glu Val Leu Thr Ser Gly Thr Val Tyr
    365                 370                 375
```

7

```
ctg gct gac att ctc ttc tgt gag tca gcc ctc ttt tat ttc tct gag    1322
Leu Ala Asp Ile Leu Phe Cys Glu Ser Ala Leu Phe Tyr Phe Ser Glu
380                 385                 390                 395

tac atg gaa aaa gag gat gca gtg aat atc tta caa ttc tgg ttg gca    1370
Tyr Met Glu Lys Glu Asp Ala Val Asn Ile Leu Gln Phe Trp Leu Ala
                    400                 405                 410

gca gat aac ttc cag tct cag ctt gct gcc aaa aag ggg caa tat gat    1418
Ala Asp Asn Phe Gln Ser Gln Leu Ala Ala Lys Lys Gly Gln Tyr Asp
                415                 420                 425

gga cag gag gca cag aat gat gcc atg att tta tat gac aag tac ttc    1466
Gly Gln Glu Ala Gln Asn Asp Ala Met Ile Leu Tyr Asp Lys Tyr Phe
            430                 435                 440

tcc ctc caa gcc aca cat cct ctt gga ttt gat gat gtt gta cga tta    1514
Ser Leu Gln Ala Thr His Pro Leu Gly Phe Asp Asp Val Val Arg Leu
            445                 450                 455

gaa att gaa tcc aat atc tgc agg gaa ggt ggg cca ctc ccc aac tgt    1562
Glu Ile Glu Ser Asn Ile Cys Arg Glu Gly Gly Pro Leu Pro Asn Cys
        460                 465                 470                 475

ttc aca act cca tta cgt cag gcc tgg aca acc atg gag aag gtc ttt    1610
Phe Thr Thr Pro Leu Arg Gln Ala Trp Thr Thr Met Glu Lys Val Phe
                480                 485                 490

ttg cct ggc ttt ctg tcc agc aat ctt tat tat aaa tat ttg aat gat    1658
Leu Pro Gly Phe Leu Ser Ser Asn Leu Tyr Tyr Lys Tyr Leu Asn Asp
            495                 500                 505

ctc atc cat tcg gtt cga gga gat gaa ttt ctg ggc ggg aac gtg tcg    1706
Leu Ile His Ser Val Arg Gly Asp Glu Phe Leu Gly Gly Asn Val Ser
        510                 515                 520

ccg act gct cct ggc tct gtt ggc cct cct gat gag tct cac cca ggg    1754
Pro Thr Ala Pro Gly Ser Val Gly Pro Pro Asp Glu Ser His Pro Gly
        525                 530                 535

agt tct gac agc tct gcg tct cag tcc agt gtg aaa aaa gcc agt att    1802
Ser Ser Asp Ser Ser Ala Ser Gln Ser Ser Val Lys Lys Ala Ser Ile
540                 545                 550                 555

aaa ata ctg aaa aat ttt gat gaa gcg ata att gtg gat gcg gca agt    1850
Lys Ile Leu Lys Asn Phe Asp Glu Ala Ile Ile Val Asp Ala Ala Ser
                560                 565                 570

ctg gat cca gaa tct tta tat caa cgg aca tat gcc ggg aag atg aca    1898
Leu Asp Pro Glu Ser Leu Tyr Gln Arg Thr Tyr Ala Gly Lys Met Thr
            575                 580                 585

ttt gga aga gtg agt gac ttg ggg caa ttc atc cgg gaa tct gag cct    1946
Phe Gly Arg Val Ser Asp Leu Gly Gln Phe Ile Arg Glu Ser Glu Pro
            590                 595                 600

gaa cct gat gta agg aaa tca aaa gga tcc atg ttc tca caa gct atg    1994
Glu Pro Asp Val Arg Lys Ser Lys Gly Ser Met Phe Ser Gln Ala Met
        605                 610                 615

aag aaa tgg gtg caa gga aat act gat gag gcc cag gaa gag cta gct    2042
Lys Lys Trp Val Gln Gly Asn Thr Asp Glu Ala Gln Glu Glu Leu Ala
620                 625                 630                 635

tgg aag att gct aaa atg ata gtc agt gac att atg cag cag gct cag    2090
Trp Lys Ile Ala Lys Met Ile Val Ser Asp Ile Met Gln Gln Ala Gln
                640                 645                 650
```

8

```
tat gat caa ccg tta gag aaa tct aca aag tta tga ctcaaaactt          2136
Tyr Asp Gln Pro Leu Glu Lys Ser Thr Lys Leu  *
              655                     660

gagataaagg aaatctgctt gtgaaaaata agagaacttt tttcccttgg ttggattctt   2196
caacacagcc aatgaaaaca gcactatatt tctgatctgt cactgttgtt tccagggaga   2256
gaatggggag acaatcctag gacttccacc ctaatgcagt tacctgtagg gcataattgg   2316
atggcacatg atgtttcaca cagtgaggag tctttaaagg ttaccaa                 2363

<210> 32
<211> 662
<212> PRT
<213> Homo Sapien

<400> 32
Met Arg Gly Ala Gly Pro Ser Pro Arg Gln Ser Pro Arg Thr Leu Arg
1               5                  10                 15
Pro Asp Pro Gly Pro Ala Met Ser Phe Phe Arg Arg Lys Val Lys Gly
              20                 25                 30
Lys Glu Gln Glu Lys Thr Ser Asp Val Lys Ser Ile Lys Ser Ala Ser Ile
         35                 40                 45
Ser Val His Ser Pro Gln Lys Ser Thr Lys Asn His Ala Leu Leu Glu
    50                 55                 60
Ala Ala Gly Pro Ser His Val Ala Ile Asn Ala Ile Ser Ala Asn Met
65                 70                 75                 80
Asp Ser Phe Ser Ser Ser Arg Thr Ala Thr Leu Leu Lys Lys Gln Pro Ser
              85                 90                 95
His Met Glu Ala His Phe Gly Asp Leu Gly Arg Ser Cys Leu Asp
              100                105                110
Tyr Gln Thr Gln Glu Thr Lys Ser Ser Leu Ser Lys Thr Leu Glu Gln
         115                120                125
Val Leu His Asp Thr Ile Val Leu Pro Tyr Phe Ile Gln Phe Met Glu
    130                135                140
Leu Arg Arg Met Glu His Leu Val Lys Phe Trp Leu Glu Ala Glu Ser
145                150                155                160
Phe His Ser Thr Thr Trp Ser Arg Ile Arg Ala His Ser Leu Asn Thr
              165                170                175
Met Lys Gln Ser Ser Leu Ala Glu Pro Val Ser Pro Ser Lys Lys His
              180                185                190
Glu Thr Thr Ala Ser Phe Leu Thr Asp Ser Leu Asp Lys Arg Leu Glu
         195                200                205
Asp Ser Gly Ser Ala Gln Leu Phe Met Thr His Ser Glu Gly Ile Asp
    210                215                220
Leu Asn Asn Arg Thr Asn Ser Thr Gln Asn His Leu Leu Leu Ser Gln
225                230                235                240
Glu Cys Asp Ser Ala His Ser Leu Arg Leu Glu Met Ala Arg Ala Gly
              245                250                255
Thr His Gln Val Ser Met Glu Thr Gln Glu Ser Ser Ser Thr Leu Thr
              260                265                270
Val Ala Ser Arg Asn Ser Pro Ala Ser Pro Leu Lys Glu Leu Ser Gly
    275                280                285
Lys Leu Met Lys Ser Ile Glu Gln Asp Ala Val Asn Thr Phe Thr Lys
    290                295                300
Tyr Ile Ser Pro Asp Ala Ala Lys Pro Ile Pro Ile Thr Glu Ala Met
305                310                315                320
Arg Asn Asp Ile Ile Ala Arg Ile Cys Gly Glu Asp Gly Gln Val Asp
              325                330                335
Pro Asn Cys Phe Val Leu Ala Gln Ser Ile Val Phe Ser Ala Met Glu
              340                345                350
Gln Glu His Phe Ser Glu Phe Leu Arg Ser His His Phe Cys Lys Tyr
         355                360                365
Gln Ile Glu Val Leu Thr Ser Gly Thr Val Tyr Leu Ala Asp Ile Leu
370                375                380
Phe Cys Glu Ser Ala Leu Phe Tyr Phe Ser Glu Tyr Met Glu Lys Glu
385                390                395                400
Asp Ala Val Asn Ile Leu Gln Phe Trp Leu Ala Ala Asp Asn Phe Gln
```

9

```
                 405                    410                    415
Ser Gln Leu Ala Ala Lys Lys Gly Gln Tyr Asp Gly Gln Glu Ala Gln
             420                    425                    430
Asn Asp Ala Met Ile Leu Tyr Asp Lys Tyr Phe Ser Leu Gln Ala Thr
         435                    440                    445
His Pro Leu Gly Phe Asp Asp Val Val Arg Leu Glu Ile Glu Ser Asn
     450                    455                    460
Ile Cys Arg Glu Gly Gly Pro Leu Pro Asn Cys Phe Thr Thr Pro Leu
465                    470                    475                    480
Arg Gln Ala Trp Thr Thr Met Glu Lys Val Phe Leu Pro Gly Phe Leu
                 485                    490                    495
Ser Ser Asn Leu Tyr Tyr Lys Tyr Leu Asn Asp Leu Ile His Ser Val
             500                    505                    510
Arg Gly Asp Glu Phe Leu Gly Gly Asn Val Ser Pro Thr Ala Pro Gly
         515                    520                    525
Ser Val Gly Pro Pro Asp Glu Ser His Pro Gly Ser Ser Asp Ser Ser
     530                    535                    540
Ala Ser Gln Ser Ser Val Lys Lys Ala Ser Ile Lys Ile Leu Lys Asn
545                    550                    555                    560
Phe Asp Glu Ala Ile Ile Val Asp Ala Ala Ser Leu Asp Pro Glu Ser
                 565                    570                    575
Leu Tyr Gln Arg Thr Tyr Ala Gly Lys Met Thr Phe Gly Arg Val Ser
             580                    585                    590
Asp Leu Gly Gln Phe Ile Arg Glu Ser Glu Pro Glu Pro Asp Val Arg
         595                    600                    605
Lys Ser Lys Gly Ser Met Phe Ser Gln Ala Met Lys Lys Trp Val Gln
     610                    615                    620
Gly Asn Thr Asp Glu Ala Gln Glu Glu Leu Ala Trp Lys Ile Ala Lys
625                    630                    635                    640
Met Ile Val Ser Asp Ile Met Gln Gln Ala Gln Tyr Asp Gln Pro Leu
             645                    650                    655
Glu Lys Ser Thr Lys Leu
             660
```

```
<210> 33
<211> 2363
<212> DNA
<213> Homo Sapien

<220>
<221> CDS
<222> (138)...(2126)
<223> AKAP-10-5

<221> allele
<222> 2073
<223> Single Nucleotide Polymorphism: A to G

<400> 33
gcggcttgtt gataatatgg cggctggagc tgcctgggca tcccgaggag gcggtggggc        60
ccactcccgg aagaagggtc cctttcgcg ctagtgcagc ggcccctctg gacccggaag       120
tccgggccgg ttgctga atg agg gga gcc ggg ccc tcc ccg cgc cag tcc       170
                     Met Arg Gly Ala Gly Pro Ser Pro Arg Gln Ser
                      1                   5                      10

ccc cgc acc ctc cgt ccc gac ccg ggc ccc gcc atg tcc ttc ttc cgg       218
Pro Arg Thr Leu Arg Pro Asp Pro Gly Pro Ala Met Ser Phe Phe Arg
             15                  20                      25

cgg aaa gtg aaa ggc aaa gaa caa gag aag acc tca gat gtg aag tcc       266
Arg Lys Val Lys Gly Lys Glu Gln Glu Lys Thr Ser Asp Val Lys Ser
         30                  35                      40

att aaa gct tca ata tcc gta cat tcc cca caa aaa agc act aaa aat       314
Ile Lys Ala Ser Ile Ser Val His Ser Pro Gln Lys Ser Thr Lys Asn
     45                  50                      55
```

10

```
cat gcc ttg ctg gag gct gca gga cca agt cat gtt gca atc aat gcc    362
His Ala Leu Leu Glu Ala Ala Gly Pro Ser His Val Ala Ile Asn Ala
 60              65              70                          75

att tct gcc aac atg gac tcc ttt tca agt agc agg aca gcc aca ctt    410
Ile Ser Ala Asn Met Asp Ser Phe Ser Ser Arg Thr Ala Thr Leu
                80              85              90

aag aag cag cca agc cac atg gag gct gct cat ttt ggt gac ctg ggc    458
Lys Lys Gln Pro Ser His Met Glu Ala Ala His Phe Gly Asp Leu Gly
             95              100             105

aga tct tgt ctg gac tac cag act caa gag acc aaa tca agc ctt tct    506
Arg Ser Cys Leu Asp Tyr Gln Thr Gln Glu Thr Lys Ser Ser Leu Ser
         110             115             120

aag acc ctt gaa caa gtc ttg cac gac att gtc ctc cct tac ttc        554
Lys Thr Leu Glu Gln Val Leu His Asp Thr Ile Val Leu Pro Tyr Phe
     125             130             135

att caa ttc atg gaa ctt cgg cga atg gag cat ttg gtg aaa ttt tgg    602
Ile Gln Phe Met Glu Leu Arg Arg Met Glu His Leu Val Lys Phe Trp
 140             145             150             155

tta gag gct gaa agt ttt cat tca aca act tgg tcg cga ata aga gca    650
Leu Glu Ala Glu Ser Phe His Ser Thr Thr Trp Ser Arg Ile Arg Ala
                160             165             170

cac agt cta aac aca atg aag cag agc tca ctg gct gag cct gtc tct    698
His Ser Leu Asn Thr Met Lys Gln Ser Ser Leu Ala Glu Pro Val Ser
         175             180             185

cca tct aaa aag cat gaa act aca gcg tct ttt tta act gat tct ctt    746
Pro Ser Lys Lys His Glu Thr Thr Ala Ser Phe Leu Thr Asp Ser Leu
     190             195             200

gat aag aga ttg gag gat tct ggc tca gca cag ttg ttt atg act cat    794
Asp Lys Arg Leu Glu Asp Ser Gly Ser Ala Gln Leu Phe Met Thr His
         205             210             215

tca gaa gga att gac ctg aat aat aga act aac agc act cag aat cac    842
Ser Glu Gly Ile Asp Leu Asn Asn Arg Thr Asn Ser Thr Gln Asn His
 220             225             230             235

ttg ctg ctt tcc cag gaa tgt gac agt gcc cat tct ctc cgt ctt gaa    890
Leu Leu Leu Ser Gln Glu Cys Asp Ser Ala His Ser Leu Arg Leu Glu
             240             245             250

atg gcc aga gca gga act cac caa gtt tcc atg gaa acc caa gaa tct    938
Met Ala Arg Ala Gly Thr His Gln Val Ser Met Glu Thr Gln Glu Ser
             255             260             265

tcc tct aca ctt aca gta gcc agt aga aat agt ccc gct tct cca cta    986
Ser Ser Thr Leu Thr Val Ala Ser Arg Asn Ser Pro Ala Ser Pro Leu
         270             275             280

aaa gaa ttg tca gga aaa cta atg aaa agt ata gaa caa gat gca gtg    1034
Lys Glu Leu Ser Gly Lys Leu Met Lys Ser Ile Glu Gln Asp Ala Val
     285             290             295

aat act ttt acc aaa tat ata tct cca gat gct gct aaa cca ata cca    1082
Asn Thr Phe Thr Lys Tyr Ile Ser Pro Asp Ala Ala Lys Pro Ile Pro
 300             305             310             315

att aca gaa gca atg aga aat gac atc ata gca agg att tgt gga gaa    1130
Ile Thr Glu Ala Met Arg Asn Asp Ile Ile Ala Arg Ile Cys Gly Glu
             320             325             330
```

11

```
gat gga cag gtg gat ccc aac tgt ttc gtt ttg gca cag tcc ata gtc    1178
Asp Gly Gln Val Asp Pro Asn Cys Phe Val Leu Ala Gln Ser Ile Val
            335                 340                 345

ttt agt gca atg gag caa gag cac ttt agt gag ttt ctg cga agt cac    1226
Phe Ser Ala Met Glu Gln Glu His Phe Ser Glu Phe Leu Arg Ser His
        350                 355                 360

cat ttc tgt aaa tac cag att gaa gtg ctg acc agt gga act gtt tac    1274
His Phe Cys Lys Tyr Gln Ile Glu Val Leu Thr Ser Gly Thr Val Tyr
    365                 370                 375

ctg gct gac att ctc ttc tgt gag tca gcc ctc ttt tat ttc tct gag    1322
Leu Ala Asp Ile Leu Phe Cys Glu Ser Ala Leu Phe Tyr Phe Ser Glu
380                 385                 390                 395

tac atg gaa aaa gag gat gca gtg aat atc tta caa ttc tgg ttg gca    1370
Tyr Met Glu Lys Glu Asp Ala Val Asn Ile Leu Gln Phe Trp Leu Ala
                400                 405                 410

gca gat aac ttc cag tct cag ctt gct gcc aaa aag ggg caa tat gat    1418
Ala Asp Asn Phe Gln Ser Gln Leu Ala Ala Lys Lys Gly Gln Tyr Asp
            415                 420                 425

gga cag gag gca cag aat gat gcc atg att tta tat gac aag tac ttc    1466
Gly Gln Glu Ala Gln Asn Asp Ala Met Ile Leu Tyr Asp Lys Tyr Phe
        430                 435                 440

tcc ctc caa gcc aca cat cct ctt gga ttt gat gat gtt gta cga tta    1514
Ser Leu Gln Ala Thr His Pro Leu Gly Phe Asp Asp Val Val Arg Leu
    445                 450                 455

gaa att gaa tcc aat atc tgc agg gaa ggt ggg cca ctc ccc aac tgt    1562
Glu Ile Glu Ser Asn Ile Cys Arg Glu Gly Gly Pro Leu Pro Asn Cys
460                 465                 470                 475

ttc aca act cca tta cgt cag gcc tgg aca acc atg gag aag gtc ttt    1610
Phe Thr Thr Pro Leu Arg Gln Ala Trp Thr Thr Met Glu Lys Val Phe
                480                 485                 490

ttg cct ggc ttt ctg tcc agc aat ctt tat tat aaa tat ttg aat gat    1658
Leu Pro Gly Phe Leu Ser Ser Asn Leu Tyr Tyr Lys Tyr Leu Asn Asp
            495                 500                 505

ctc atc cat tcg gtt cga gga gat gaa ttt ctg ggc ggg aac gtg tcg    1706
Leu Ile His Ser Val Arg Gly Asp Glu Phe Leu Gly Gly Asn Val Ser
        510                 515                 520

ccg act gct cct ggc tct gtt ggc cct cct gat gag tct cac cca ggg    1754
Pro Thr Ala Pro Gly Ser Val Gly Pro Pro Asp Glu Ser His Pro Gly
    525                 530                 535

agt tct gac agc tct gcg tct cag tcc agt gtg aaa aaa gcc agt att    1802
Ser Ser Asp Ser Ser Ala Ser Gln Ser Ser Val Lys Lys Ala Ser Ile
540                 545                 550                 555

aaa ata ctg aaa aat ttt gat gaa gcg ata att gtg gat gcg gca agt    1850
Lys Ile Leu Lys Asn Phe Asp Glu Ala Ile Ile Val Asp Ala Ala Ser
                560                 565                 570

ctg gat cca gaa tct tta tat caa cgg aca tat gcc ggg aag atg aca    1898
Leu Asp Pro Glu Ser Leu Tyr Gln Arg Thr Tyr Ala Gly Lys Met Thr
            575                 580                 585

ttt gga aga gtg agt gac ttg ggg caa ttc atc cgg gaa tct gag cct    1946
Phe Gly Arg Val Ser Asp Leu Gly Gln Phe Ile Arg Glu Ser Glu Pro
```

12

```
gaa cct gat gta agg aaa tca aaa gga tcc atg ttc tca caa gct atg    1994
Glu Pro Asp Val Arg Lys Ser Lys Gly Ser Met Phe Ser Gln Ala Met
    605             610             615

aag aaa tgg gtg caa gga aat act gat gag gcc cag gaa gag cta gct    2042
Lys Lys Trp Val Gln Gly Asn Thr Asp Glu Ala Gln Glu Glu Leu Ala
620             625             630                         635

tgg aag att gct aaa atg ata gtc agt gac gtt atg cag cag gct cag    2090
Trp Lys Ile Ala Lys Met Ile Val Ser Asp Val Met Gln Gln Ala Gln
            640             645             650

tat gat caa ccg tta gag aaa tct aca aag tta tga ctcaaaactt        2136
Tyr Asp Gln Pro Leu Glu Lys Ser Thr Lys Leu  *
            655             660

gagataaagg aaatctgctt gtgaaaaata agagaacttt tttccttgg ttggattctt   2196
caacacagcc aatgaaaaca gcactatatt tctgatctgt cactgttgtt tccagggaga   2256
gaatgggatg acaatcctag gacttccacc ctaatgcagt tacctgtagg gcataattgg   2316
atggcacatg atgtttcaca cagtgaggag tctttaaagg ttaccaa                 2363
```

```
<210> 34
<211> 662
<212> PRT
<213> Homo Sapien
```

```
<400> 34
Met Arg Gly Ala Gly Pro Ser Pro Arg Gln Ser Pro Arg Thr Leu Arg
1               5                   10                  15
Pro Asp Pro Gly Pro Ala Met Ser Phe Phe Arg Arg Lys Val Lys Gly
                20                  25                  30
Lys Glu Gln Glu Lys Thr Ser Asp Val Lys Ser Ile Lys Ala Ser Ile
            35                  40                  45
Ser Val His Ser Pro Gln Lys Ser Thr Lys Asn His Ala Leu Leu Glu
        50                  55                  60
Ala Ala Gly Pro Ser His Val Ala Ile Asn Ala Ile Ser Ala Asn Met
65                  70                  75                  80
Asp Ser Phe Ser Ser Ser Arg Thr Ala Thr Leu Lys Lys Gln Pro Ser
                85                  90                  95
His Met Glu Ala Ala His Phe Gly Asp Leu Gly Arg Ser Cys Leu Asp
            100                 105                 110
Tyr Gln Thr Gln Glu Thr Lys Ser Ser Leu Ser Lys Thr Leu Glu Gln
        115                 120                 125
Val Leu His Asp Thr Ile Val Leu Pro Tyr Phe Ile Gln Phe Met Glu
    130                 135                 140
Leu Arg Arg Met Glu Leu His Leu Val Lys Phe Trp Leu Glu Ala Glu Ser
145                 150                 155                 160
Phe His Ser Thr Thr Trp Ser Arg Ile Arg Ala His Ser Leu Asn Thr
            165                 170                 175
Met Lys Gln Ser Ser Leu Ala Glu Pro Val Ser Pro Ser Lys Lys His
            180                 185                 190
Glu Thr Thr Ala Ser Phe Leu Thr Asp Ser Leu Asp Lys Arg Leu Glu
        195                 200                 205
Asp Ser Gly Ser Ala Gln Leu Phe Met Thr His Ser Glu Gly Ile Asp
    210                 215                 220
Leu Asn Asn Arg Thr Asn Ser Thr Gln Asn His Leu Leu Leu Ser Gln
225                 230                 235                 240
Glu Cys Asp Ser Ala His Ser Leu Arg Leu Glu Met Ala Ala Ala Gly
                245                 250                 255
Thr His Gln Val Ser Met Glu Thr Gln Glu Ser Ser Ser Thr Leu Thr
            260                 265                 270
Val Ala Ser Arg Asn Ser Pro Ala Ser Pro Leu Lys Glu Leu Ser Gly
            275                 280                 285
Lys Leu Met Lys Ser Ile Glu Gln Asp Ala Val Asn Thr Phe Thr Lys
        290                 295                 300
```

13

```
Tyr Ile Ser Pro Asp Ala Ala Lys Pro Ile Pro Ile Thr Glu Ala Met
305                 310                 315                 320
Arg Asn Asp Ile Ile Ala Arg Ile Cys Gly Glu Asp Gly Gln Val Asp
                325                 330                 335
Pro Asn Cys Phe Val Leu Ala Gln Ser Ile Val Phe Ser Ala Met Glu
            340                 345                 350
Gln Glu His Phe Ser Glu Phe Leu Arg Ser His His Phe Cys Lys Tyr
        355                 360                 365
Gln Ile Glu Val Leu Thr Ser Gly Thr Val Tyr Leu Ala Asp Ile Leu
    370                 375                 380
Phe Cys Glu Ser Ala Leu Phe Tyr Phe Ser Glu Tyr Met Glu Lys Glu
385                 390                 395                 400
Asp Ala Val Asn Ile Leu Gln Phe Trp Leu Ala Ala Asp Asn Phe Gln
                405                 410                 415
Ser Gln Leu Ala Ala Lys Lys Gly Leu Tyr Asp Gly Gln Glu Ala Gln
            420                 425                 430
Asn Asp Ala Met Ile Leu Tyr Asp Lys Tyr Phe Ser Leu Gln Ala Thr
        435                 440                 445
His Pro Leu Gly Phe Asp Asp Val Val Arg Leu Glu Ile Glu Ser Asn
    450                 455                 460
Ile Cys Arg Glu Gly Gly Pro Leu Pro Asn Cys Phe Thr Thr Pro Leu
465                 470                 475                 480
Arg Gln Ala Trp Thr Thr Met Glu Lys Val Phe Leu Pro Gly Phe Leu
                485                 490                 495
Ser Ser Asn Leu Tyr Tyr Lys Tyr Leu Asn Asp Leu Ile Ile Ser Val
            500                 505                 510
Arg Gly Asp Glu Phe Leu Gly Gly Asn Val Ser Pro Thr Ala Pro Gly
        515                 520                 525
Ser Val Gly Pro Pro Asp Glu Ser His Pro Gly Ser Ser Asp Ser Ser
    530                 535                 540
Ala Ser Gln Ser Ser Val Lys Lys Ala Ser Ile Lys Ile Leu Lys Asn
545                 550                 555                 560
Phe Asp Glu Ala Ile Ile Val Asp Ala Ala Ser Leu Asp Pro Glu Ser
                565                 570                 575
Leu Tyr Gln Arg Thr Tyr Ala Gly Lys Met Thr Phe Gly Arg Val Ser
            580                 585                 590
Asp Leu Gly Gln Phe Ile Arg Glu Ser Glu Pro Glu Pro Asp Val Arg
        595                 600                 605
Lys Ser Lys Gly Ser Met Phe Ser Gln Ala Met Lys Lys Trp Val Gln
    610                 615                 620
Gly Asn Thr Asp Glu Ala Gln Glu Glu Leu Ala Trp Lys Ile Ala Lys
625                 630                 635                 640
Met Ile Val Ser Asp Val Met Gln Gln Ala Gln Tyr Asp Gln Pro Leu
                645                 650                 655
Glu Lys Ser Thr Lys Leu
                660
```

```
gaattcctat ttcaaaagaa acaaatgggc caagtatggt ggctcatacc tgtaatccca    60
gcactttggg aggccgaggt gagtgggtca cttgaggtca ggagttccag gccagtctgg   120
ccaacatggt gaaacactgt ctctactaaa aatacaaaaa ttagccgggc gtggtggcgg   180
gcacctgtaa tcccagctac tcaggaggct gaggcaggag aattgcttga acctgggaga   240
tggaggttgc agtgagccga gatcgcgcca ctgctctcca gcctgggtgg cagagtgaga   300
ctctgtctca aaaagaaaca aagaaataaa tgaaacaatt ttgttcacat atatttcaca   360
aatttgaaat gttaaaggta ttatggtcac tgatatcctg tttcattctt tatataatca   420
ttaagtttga aatgtatact tgcactacta acacagtagt taatcttagt cctacaagtt   480
actgcttttа cacaatatat tttcgtaata tgtatgcact ggtgtttatg tacgtgttta   540
tgtttatatc tgttaaaatt agcagtttcc atctttttct attttgtacc atcacatcag   600
```

```
ttcagaagga ttgacagagc aaaatgattt gatgaagtat aaaagtcaca tggtgagtgg      660
cataaataca actctgaaca attaggaggc tcactattga ctggaactaa actgcaagcc      720
agaaagacac atatcctata tgtcaagaga tgtaccaccc aggcagttaa agaagggaag      780
tacacataga aagcacaatg gtgaataatt aaaaaattgg aatttatcag acactggatt      840
catttgctcc taaagtcaga gtcctctatt gtttttttgt ttttgtgggt ttcttttaa      900
atttttttat tttttgtaga gtcggagtct caatgtgtta cccgggctgg tctagaactc      960
ctggcctcaa acaaacctcc tgcctcagct tcccaaagca ttgggattac agacatgagc     1020
cactgagccc agcccagacg ctttagcatt tgaatgcatt ctgaaatagt tgtagaaacc     1080
gcataagctt tccatgtcac tttcaaagtt tgatggtctc tttagtaaac caaccaagtt     1140
attcctcaag ggcaaaataa catttctcag tgcaaaactg atgcacttca ttaccaaagg     1200
gaaaagacca caactataga ggcgtcattg aaagctgcac tcttcaggag ccaaaaaaa     1260
aggtacaaac acatactaat ggaacattct ttagaagagc cccaaagtta atgataaaca     1320
ttttcatcaa aagaaaaaga gaacaaggtg ttagcaaatt cctctatcaa ataacactaa     1380
acatcaagga acatcaatgg catgccatgt ggaagaggaa gtgctagctc atgtacaaac     1440
cagtagataa tttcaacttg ctgccgaatg aaacctcttt gcaaggtatg aatcagcact     1500
tctcatgttt gttttgcttt gttttgtttt gttttagag acaggccctt gctctgtcac     1560
acaggctgga gtgcagtggc acgatcagag ctcactgcaa cctgaaacct ctgggctcaa     1620
gggatcctcc tgccttagcc tcccaagtag ctgggactac aggcccacca tgcccagcta     1680
attttttaaa ttttctatag agatgggatt tcactagcac ctttcatgtt tgatgttcat     1740
atacaacgac caaggtacaa tgtggaaaag ggtctcaggg atctaaagtg aaggaggacc     1800
agaaagaaaa ggggttgcta catagagtag aagaagttgc acttcatgcc agtctacaac     1860
actgctgttt tcctcagagc agagttgatg atctaaatca ggggtcccca accccccagtt     1920
catagctgtt taggaaccgg gccacacagc aggaggtgag caataggcaa gcgagcatta     1980
ccacctgggc ttcacctccc gtcagatcag tgatgtcatt agattctcat aggaccatga     2040
accctattgt gaactgagca tgcaaggagt gtaggttttc cgctcttat gagactctaa     2100
tgccggaaga tctgtcactg tcttccatca ccctgagatg ggaacatcta gttgcaggaa     2160
aacaacctca gggctcccat tgattctata ttacagtgag ttgtatcatt atttcattct     2220
atattacaat gtaataataa tagaaataaa ggcacaatag gccaggcgtg gtggctcaca     2280
cctgtaatcc cagcacttcg ggaggccaag gcagcggat cacgaggtca ggagatcgag     2340
accatcctgg ctaaaacggt gaaaccccgt ctactaaaaa ttcaaaaaa aattagccgg     2400
gtgtggtggt gggcacctgt agtcccagct actcgagagg ctgaggcagg agaatggtgt     2460
gaacctggga ggcagagctt gaggtaagcc gagatcacgc cactgcactc cagcctgggc     2520
gacagacgga tactctgtct caaaaaaaa aaaaaaaaa aaagaaataa agtgaacaat     2580
aaatgtaatg tggctgaatc attccaaaac aatcccccca ccccagttca cggaaaaatt     2640
ctcccacaaa accagtccct ggtgccaaaa aggttgggga ccgctaatct aaataatcta     2700
atcttcattc aatgctaaaa aatgaataaa cttttttta aatacacggt ctcactttgt     2760
tgcccaggct ggagtacggt ggcatgatca cagctcactg tagcctcaat caaccaggcc     2820
ccagctaat ttttaaattt tttatagaga tggggtctc accatgttgc ccagactggt     2880
ctcaaaccct gggctcaagt gatcctccct caaactcctg gactcaagtg atcctccttc     2940
cttggcctcc caaagtgctg ggattacaag catgagccac tgtacccagc tggacagtga     3000
ttttaagtcg cactacagtc atggacaatc aggctttca acatgcagta tggacagtga     3060
gtcccagggt ctgctttttcc atactgaaat acatgtgata ctaaggagaa aggtgctcgc     3120
aaggatattt aaaatgaaga atatttaaaa tgaggaaaaa actgtttctt catgactttg     3180
ataaggctga taaagaccat ttctgtgatc tcaggtgatt cactcaagta gtatatttca     3240
gtaatcatta tctggaacag cctgaatctt aaccaaaata ccatgatttt ttaatgctgt     3300
tatgatacct tgatgatatg accaaactgc aatgtaggcca gctaaatctc cacgagtttg     3360
acttccccga gagttgacag tttcttcac aatataatttt ttgatacatg     3420
attggcataat ttaaaaacta cactgaaatg ctgcaaaatg atataaagaa acattttcca     3480
gaatcaaatg caatcaaaga gtggattaag aatctatca ccattatcaa ctaaatagaa     3540
acacttggac tgggtgtggt ggctcacacatc tgtaatctca gcactttggg aggccaaggc     3600
aggtggattg cttgaggcca ggagctcaag accagcctga gcaacatagc aaaactctgt     3660
ctctacaaaa aaaaaaaaaa aaaatcacagg catgtggca gatgcttgta atcccagcta     3720
ctctgaagc tgaagtagga ggatcgcttg agcccaggag atcaagactg cagtgagccg     3780
tggtcatgct ggcccacagc ctgagtgaca gagagagacc ctgtctcaaa aacaaaaaca     3840
aacaaaaaac acttaacctt cctgtttttt gctgttgttg ttgttgtttg tttgttttga     3900
gatggagtct cactctgttg cccaggctgg agtgcagtgg cgtgatcttg gctcactgca     3960
agctctgcct cccgggttca cgccattctc ctgcctcage ctcccgagta gctgggacta     4020
taggcgcccg ccaccacgcc cggctacttt tttgcattttt tagtagagat ggggtttcac     4080
cgtgttagcc aggatggtct tgatctcctg acctcgtgat ccacccgcct cggcctccca     4140
aagtgctgag attacaggca tgagccaccg caccggccca acctttctgt tttttagttt     4200
gatatgcttg ttaacttcagc agctgaaaga atgctgaaag tggccttcag taaaaaaatt     4260
tcactagaat ctctacatcc atatttaatc tgaatgcata tccagattga tcagttagag     4380
caaaaacaat catcatcatt cctgatgacc tctaattctg gtttcggctt tctattttcaa     4440
tggaaacaga ataaggaaag aaatggaagg gtctctggaaa tttgtcctgg gctatagtata     4500
ctatcaaaga tcaccaacaa taagatctct cctataaata taaaacaagt ataattaatt     4560
ttttaattat tttttttctct tcagaggatt ttattccaag ataaaacata acttctaccc     4620
```

                                    15

```
atactattga ttccaaaggt tagaaaaagt gttttttcctc atcttatcct tcaaagaggt    4680
cacagcaatg caaacatcta taaaatgcct ctgcataatt gtcagaagct atagtccaga    4740
aatcattgaa aatgcttttc cattttaagc ttaggtgagg tgtcttagga aacctctatg    4800
acaacttact ctatttattg ggaggtaaac tcccagactc tcccagggtc tcctgtattg    4860
atctcatttt ttaggcttcc taatccttg aagcacaatc gaaaaagcc tggatctctt     4920
ttctgcacat atcatcgcgg aattcattcg gcttccagca agctgacact ccatgataca    4980
agcggcctcg cccttctccg gacgccagtc cttgctgcgg ttagctagga tgaggggttt    5040
gctgggcttc agtgcaggct tctgcgggtt cccaagccgc accaggtggc ctcacaggct    5100
ggatgtcacc attgcacact gagctcctgg caggctgtac caattttta attatttaat     5160
atttatttt aaaattatgg tgaatatttt ggtatttctgc tctaaaatag gcccataaat    5220
gcacagcaga tatctcttgg aacccacagc tttccactgg aagaactaag tatttttctt    5280
ttaaagatgc tactaagtct ctgaaaagtc cagatcctct acctctttcc gcccaaaact    5340
aagacttgga atttatgaga gatctagcta acagaaatcc cagacacatc attggttctt     5400
cccagagtgc agtcctccta aagaggctca ggcctaagca ggccctgca ccaggagggt     5460
gggtctgaga cccacatagc acttcccaag gtgcatgctc cagagaggca ctgaaacagc    5520
tgagcacaag cctgcaagcc tggagaactc tcacagtcag aacggagggg gcccagtggg    5580
actaacataa agagaaaagg gaacacagag aaatggatgg caccaacaac cagcaaagcc    5640
ttcatggcca atgaaaagcat cagtgacggg gccagaaccc tcatccccaa agactcttca    5700
ctgcctttag tgaaaaacaa tggctagaga gtgaagttat gatcatgtat agacaggtaa    5760
agttacattt ttatattctg actctgctaa tgtgaaattc cctatctgct agactaaaag    5820
tttcagacac cctgttcaaa tatcccatta gttgctagag acttaaaatg aacagaacgc    5880
acattgttcag gatgactatt accaaaaaat caaaagacag caagtattgg tgggatgta    5940
gagaaactg aactttttgtg cactgtttat gagaatgtaa aatggagcag ctgctgtgga     6000
aaagatgatg caggttcctc aaagagtaaa accaagatgt ggaaacaact aaatgcccat    6060
cagtggatga aggggtagac aatatgtgt atatacatac catggagtac tattcagcct    6120
ctaaaaaaa aaaaggaaat tctataacat gcaacagcat ggatgaatct tgaggacatt    6180
ttgctaatga aataaggcag tcatagaaag acaaatactg cacgactcca cttatatgaa    6240
ataccaaaaa tagacaaatt catagaatca aagagtacaa tggaggttac ctggagctgc    6300
agggcgggaa acgagagtt actaatcaac gaacataacg ttgcagttaa gtaagatgaa    6360
taagctctca agatcagctg tacaacactg tacctagagt caacaataat gtattgtaca    6420
cttaaaaatt tgttaagggt agattaacaa atgtagtaga tccacaaatg tggttaagtg    6480
ttcttaccac agtaaaaataa aaaaagaata tcaagcccag gagttcgaga ctagcctggg    6540
taacatggtga aaaccctgtc tctacagaaa atacaaaaat tagccagctg tggagggtgca    6600
ctcctaggga ggctgaggtg ggaggcttgc ttgagcccag gaggtcaagg ctgcagtgag    6660
ccatgattgc accactgtac tccagcccag atgacagagc aagacaccac ccccccaaa    6720
aaaagaaaaa gaatatcaaa cattttaaaa gatcagatac gcaagaacaa caacaaaaaa    6780
gagatgaaca gagcatcgac cctcatctag tgggattctt ggtctaactg aaaaacagac    6840
attgagagac aaacaatgac agtgatgtga tcacagcaat tacacagtta tccctgggg    6900
actgcagaag aaaggaggaa tgcctaactt tcagaaaata gagaaagcgt caaacagtga    6960
gtgaaagcct tccaaaacta gagagaactg cacacaccaa atcacagaaa gaagaaaagc    7020
cgtgggagat tctgggaccc accggctatt tttgatggct gaaccccctg ctgcaggaga    7080
gacaggagct ggaaagcatg gtgggatgaa acctcaaaca gcttgcctg cattgcttaa    7140
gatgactggg cttgattaac tctagtcaat gggcagcaatt caatcaaaga agaaagatgc    7200
tcaaattcac attttagaat gatttttttat ggcagtatgg ggaatagatt aaaagagagt    7260
gaagctggag gcaagaaact tgttaaggag caactgaaac agtctagatg ataaataata    7320
aactgacaga gtgactagaa aaatcagaac aggctgaatc aacagatacc tagatgaaaa    7380
taacaggact tgatcaccag ttgtatcttg gagaggaagg agttgtttcc ttgctttccc    7440
tacgactggg aatacggaag gtttgccgtg tgtattggtt atatactggt gtgtagccaa    7500
tcactgacaa ccatttagca gcttaaaaca caaaggctta tctcccagtt tctgtgggcc    7560
aggaatctaa gataggctta gctggctggt tctggctcaa agtttctcaa gaggttgcaa    7620
tcaagatgtc agctgggggtt gcatcatctg aaggctcaac tggggccgga gggtccactt    7680
ccaaggagtt cactcacctg cctgacaagg cagtgctggt tgttggcagg agatctcaat    7740
tcattgccaa gtgagcctct ctatagcatt gctggaacat cctccccatc tggcagttgg    7800
cttctctcaa catgagtgat ctgagagaga gagcaaggag gaagccacag tgttcttcct    7860
actcctactc ctaacactat ggacctactc ctaacactct cacttctgcc ttattccatt    7920
agttagaaag ggaactaagc tccacctctt gaaataagaa gtgtcaaaga atttgttggat    7980
atatttaaaa atcatcacac tgtggaagtg gatagggggt tcaattaatg ctgaacttga    8040
aatgcctgga acattcaaat gtccaacagg caatgaacat acccatagat ggtcatgact    8100
ttagcaagaa tagaggaaga tcacagaatt aaggaggaat tgaaaggtaa aagaagtgga    8160
gtcagattcc ccctgaaaag tgagccatga aaggaacttt aactattgag ttagaggtca    8220
gagtaggaaa tttcggttgga attctttttt aaagaaaaga accatataag catgtttga    8280
ggtagaggga gaataaatca gtagacaggg agaggtaaaa aacataaatg ataggggtaa    8340
gttgacaaag gtcttggcag aatccttac ccattgactt ggggccaaga gagggacact    8400
tcttttgtttg agggataagg aaaataagaa agaatgggtg ctatttagtg tggtcctgtc    8460
tctagggcaa acgcataggt aacaaactgt gtgtgttagg aatatagatg tgacctcaca    8520
ttgagattct cacctcaaat ccattttgtt gttacctgta ccttcctacc ttctcttttt    8580
gctacatgca gactgctgtt ttgtcttcct ggcctgttcc aggtttcagc attctggcat    8640
```

                                    16

```
atctgctacc ctgttcccaa acctctctag agtccatgct ccttccttgg atagtgtttg    8700
attgggccac gtatctaaga agtgatgcct tcagttaggc ctgagaacct cctctatgga    8760
aatctccatc agtgaccctg acagacttgg tatcttggag atgtcactgc tcccagcctg    8820
tggtctagga gaatctcagc ctgggcctct agtagtatgg ataaggcgtt aaggtatctt    8880
tgaaccagag tctgtcatat tcctcaatgt gggacagata aaacagtggt agtgctggtg    8940
tttctgagct agaactctgg tttttggtct agattctttg atgtatgacc tttcagaggt    9000
attaaaatt  gttctaatac aatgttcaat acaaatgtag ttcctttct gttaggacct    9060
caacaaaaca tgaccaactg tagatgaaca ttaaactatg acaattcatg gaagatgaata  9120
cagtaatacc tgcggttccc ccattttagc agtcactatg gtgacatttg gcacaaatgg    9180
ctatttaagg gtgcttttgt taaaacctac catcttacta ggcacatgat attgaaacta    9240
atgaaataat ggagaaactt cttaaaaact tttaatgaat aaagtgatga agtgataata    9300
ttttagctgc tatttataaa gtgactatta caggtcaaac attcttctag ggttttttg    9360
ttgaagttgt cacatttaat ccttaataac ccactatgag tcaggtattc ttctctcccc    9420
tttggacagt tggggaaatg gggtcagag  aggttaggta atttgctcag ggccacacaa    9480
cctgcatgta gaaaatctga gatttgttaca ggaacgtatc aaactctgaa gtccatgctt   9540
ctattttcc  atgctgcctt tctaataaaa ggtaactaat gctactgagt gctgcccca    9600
aagtgagtca ctttcacccc accctacttg attttctcca taaaactaat caacactatt   9660
caacttattt attgctgatc tcccccacta gattataaac tcaataaaag caagatcctt    9720
gtctgctgaa tatcagtacc taaaacgctg tctagcacag agcaagtaat taatatttgt    9780
tgaatgaaca aataaaggaa aaaaattcaa aggaagaaaa agccctaaaa cagatgttta    9840
cctaaacata cattttaaaa gaaagcatat aaacaaattca ggacagaatt taaatttgat   9900
tttttaaaga aataaccaag tgctagctgg gcacagtggc tcacacctgt aatcctagca    9960
ctctgggaag ccgaggcagg cagatcactt gaggtcaaga gttcaagacc agcctggcca    10020
acatggtgaa acctgtctct actaaaaata cagaaattat ccaggcatgg tggcaggtcc    10080
ctgtaacccc agctactcag gaggctgagt caggagaatt gcttgaaccc aggaggcaga    10140
ggttgcagtg ggccaagatt gcaccactgc actccagcct gagtaacaaa gcaagactct    10200
gtctgaagga gaagaaaga  aagaaggaaa gaaggaaaga aggaaagaag gaaagaaaga    10260
aagaaagaaa gaaagaaaga aagaaagaaa accaagtgct tatttgggac ttactatgct    10320
aagaaagaaa gaaagaaaga gaaagaaaga accaagtgct tatttgggac ttactatgct    10380
atgtttttcc atgcacgcta ttttcagtaa agcagttagc aaacttgcaa gatcataaca    10440
acaaatatat gcttctataa ctctaaaatt gtgctttaag aagttcctct ttaccagctc    10500
atgtatgcat tagtttcta  agagttacta gtaactttt  ccctggaaga tatccacagc    10560
cagtttattt aaccaaagga ggatgcttac taacatgaag ttatcaaatg tgagcctaag    10620
ttgggccagt tcatgttaat atactccaga acaaaaacca tcctactgtc ctctgacaat    10680
tttacctgaa aattcatttt ccacattacc aaggagccag ggtaggagaa tatagaaaga    10740
ccacccaaga atccttactt cttcagcaa  aatcaattca aagtaggtaa ctaaacacat    10800
gccctaacaa tgaatagcag attgtgctca gaagaatgat ctacaacatc ttactgtgaa    10860
ggaactactg aaatattcca ataagacttc tctccaaaat gatttttatg aatttgcatt    10920
ttaaaaaata tttaagcct  aaattttaaa aggtttgata ttggtacatg aatagacaaa    10980
cagacatgga ctagaccaag aattaggttc aaacatatac aggaatttaa tatacgataa    11040
atctagtatt ccaaaggaac caacaaatgg tgttcagaca gcaggatagg cataacacat    11100
aacacagttg ggcaccctac cttactccta acaccaggag taactgaagg agcaccaaat    11160
atttatttat tttaattata gtttaagtt  ctaggtgtacg tgtgcacaac atgcaggttt    11220
attacatagg tatacatgtg ccatgttggt gaggagcacc aaataatttaa aagaaaaaaa   11280
ttggccaggg gcggtggctc acacctgtaa tcccagcact tggggaggcc aaggtgggca    11340
gatcacctga ggtcgggagt tcgagaccag cctgagcaac atggagaaac cccatctcta   11400
ctaaaaataac aaaattagcc aggcatggtg gcacatgcct gtaatcccag ctacttggga   11460
ggctgaggca ggagaatcgc tttaatctgg gaggccaggg ttgcggtgag ctatgatcca   11520
gcactccagc ctgggcaaca agagcaaaac ttcaactcaa aaaaattaat aaataaataa    11580
aaataaagaa agaaaagaaa aaaatgaaaa tagtataatt agcagaagaa aacaccgtag   11640
aatcctcgga ctcttaggat ggggaatgcc tataatataa aaaccctgaa gttataaaag    11700
agaaaatcac ctacatacaa accaaatctt tctacatgcc taaaacatag cacaaacaca   11760
gctaaaataat catagctgaa tgaactggga aaacaaaact tgactcatat ccagacagag  11820
ttaattttcc tacacataaa gagtacctat ataaacccaa caaaaaaacc accactaacc    11880
caaaataaaa atgtgacagg taatgaacag gtagttcaca gagaatacaa atggctcttc   11940
ggcacataag atgctcagac tgactttac  ttattattt  tttgagagac agggtctcac   12000
gatgttgccc aggttaggct caaactcctg ggctcaaatg atagtaccag gactacaggt   12060
gtgccccacc gcacctggct cctcaaccac ctgtattaac aggaaatgca aaataactc    12120
ttcaaatcta ttttacctat tagaatggca aaaatttgaa aaacttcaaa catcatcatg    12180
ttggtgagaa tgtgagagga ctggcactct cattttttgc tgatagcata tatactactga  12240
tggcttctat ggaaagcaat ctggcagcgt ctatcaaatg tacaagtgca tatatccttt   12300
gacaaagcaa ttccactcta ggaatgtgtt ctatatggtt gtgcttcctg gggctgggaa   12360
ctgggagcta agggacaggg gcagaagata atcttcttt  ccctccttcc ccgttaaaca   12420
tgttgaattt tatatactgt aatatattat ttttcacaaa agataatttt taagcgataat  12480
gtctgggaat tttttttttt ctttttctgag acagggtctc actctgtcat ccaggctgga  12540
atgccatggt atgatctcag ctgactgcag cctcgacctc ctgggttcaa gcaatcctcc   12600
cacctcagcc tcctgagtag ctgggactac aggcacgtgc catcatgtca attttgtgat   12660
```

17

```
atacagggtc tcactatgtt gcccaggcta atgtcaaact cctaggctca agcaatccac   12720
ccacctcagg ctccaaagtg ctgggattac aggcgtgagc caccgcgcct ggccctggga   12780
attcttacaa aagaaaaaat atctactctc cccttctatt aaagtcaaaa cagagagga    12840
aattcaacct ataatgaaag tagagaaggg cctcaaccct gagcaacaaa cacaaaggct   12900
atttctgaga caggaatttg ctgaacaaaa tcgagggaag atgacaagaa tcaagactca   12960
cttctcggct gggcgcagtg gctcacacct gtaatcccag cacttcggga ggccgaggcg   13020
gacagatcac gaggtcagga gattgagacc atactggcta acacagtgaa acccagtctc   13080
tactaaaaat aaaaaaatt agccgggcgt ggtggcagtc gcctgtagtc ccagctactt    13140
gggaagctga ggcaggagaa tggcgtgaac ccaggaagcg gagctgcag tgagccgaga    13200
tcacgccact gcactccagc ctgggtgaca gagcaagact ctgtctcaaa aaaaaaaaa    13260
aagactcatt tctctagatc ttgagccgta tccaaattta tctcagctta gtgagaggtt    13320
aaagcaagga atatccttcc ctgtgggccc tgctccttac tgaaggaagg taacggatga    13380
gtcaaggaca ccaatggaga aaagcactaa caccattatc tgatgaacat tacgtgaaga    13440
aaatggaata taggaaatcc ataaaagtga ttaaaaagat gttaaggcat gaggcgggtg    13500
gaccacaggg tcaggagatc gagaccatcc tggctaacac ggtgaaaccc catctctact    13620
aaaaatacaa aaaattagcc agcgtggtg gcaggcacct gtagtcccaa ctactcggga    13680
gactgaggca ggagaatggc atgaacctgg gagacggagc ttgcagtgag ccagatcac    13740
gccactgcac tccagcctgg gtgacagagt gagactccat ctcaaaaaaa aaagttagat    13800
acgagagata aagatccaac agacacacaa ctgctaattc tgaacagaac aaaacaaatg    13860
gcacaggaaa agaaaattta agatataaca ccggaaaact ttcctgaaat tgagtaactg    13920
aatctatagc ttgaaagggt ttagcatatg ccaagaaaaa tcagtagagt ccaaccagca    13980
caagacacat ctagcaaggc tggtgattct accaacacag agaaagaagt gggtgaccca    14040
taatgcggaa aaaggcagac catctgcagt cttctccaga acactgaagt ctgaagacaa    14100
aagaatgctg cctactgagc cagaagggag agaaagtgac ccaacacatc tttaccaagt    14160
tagaatgtca cgcattattt aaaggctgca aaagccatga aagacatgaa agaacacaag    14220
catttacaac atgaaagaac acaagcattc tcatactcaa gaatccttaa gaaaaatgta    14280
gtcctaatcc agcccactga aagttaaatg tacttaatgt gctcattaat gggaacttca    14340
tagcttcaaa tcagtctggt cccatctacc aacatctctc gcccggcttt cctgcaatag    14400
tcagcacctt tccctcctcc cagtcttgtc ccctggagtc tgctctcagc atagcagagt    14460
gaccacatca acaccaagt cagagccctc cagtgcgcac tggtctacaa agccctttccc    14520
acccccccacc ccacgtgccc tccggatcct tgtgacgtgt ctcctgcata ccctagcagc    14580
cctggcctcc tcactgcccc tctgtacat caggaaggcg actccttgag tcttggctct    14640
ggccgcctcc tccacctgca gtgagttaac tcccttacct actctaggtc attgctcaaa    14700
tgtcagcatc tcaatggggc cctccctgac taccctattt aaattctaca tactccccctt    14760
gacccatgg acctcactca ccctattcca cttttattct tacaatttag cacttgttct    14820
cttctaacgt attctaagac ttactcattt attacattgt ttgccacccc ctctagtaca    14880
taaactccag aggggcaggg atttctgtct atttattcat ttctttatcc ctaggacata    14940
gaacagggca tagttcagag tattcaatgt tatcaatgaa tgaactagca gtagtaccag    15000
ttccagttag gcacagaatt aaatctaaat agaattaaat ctcatggtct gggttaacta    15060
tggatagaaa attagatata attttaagaa gcctagaaag aaaaaattaa taatgtaaaa    15120
ataatattaa tttgataata ataacaaaaa ctctgccagg cactgtggct caaatctgca    15180
atcccagcta ctcagggagc tgaggtggaa ggatcacttg agaccagagt tcaagactca    15240
gcctaggcaa cacggcaaga aactgtctct aaaaaaaata aaacttaaat ttttctcaagt    15300
gaattctcaa agcgtcacaa aaactggaga ttaagtaaca ggaagtgtga agtaatatta    15360
ctatgctaat ggtttttttt tttttagaa aggataaacc aaaagatttc tttctcaagt    15420
cgataaactg agaaagataa gcatatcttc caattaacag aggggaggga aaagccagat    15480
acaacaaaat aagatataaa ttagttctca gttgaaaaca agagtaggag ttattttgca    15540
tcacctcaac tgtgacctcc cccagcccaa aaaacactac tgataaacag ggtagaaaag    15600
catcatctca gataaagcag gaaaactgc cacagtctca aaccacaaac tataagcaca    15660
cacctgggca accctgccaa gtctgggctc agtaggagga acgtgctgag agctaggatg    15720
taccaactta gacattctgt gggatacaga tgtccctgga agggtcacac catctcaaag    15780
gcacctgga tgcccactga tcacagccac catatgtgag agagaaactc agggcactta    15840
gagagtataa caagaacctt atgtcatctg agatgaggaa tcctcagccc tgcaaattaa    15900
ccaactcttt agaacaactg gcaaaacata aatatccaca acttttgttt cagtaattcc    15960
actctttagat atcaatcaa agtacatgag acagcagata cacacacaaa atggtatttca   16020
ctgcagcatt gtttataata gcaaaaaaca agaataatc catatgtctc aataggatac    16080
tgggtacatg agggtatgta cccatcattc aaccatcaaa aagagtgata tggatggcca   16140
cagatggaca taaaaagctg tgtgttacgt gaaaacaaac tcaagcagca gcaggatggg   16200
cttatgatag tcagtatgag ctaatttctg gaaaaaaaaa tctagtgtgt gcacagaaaa   16260
catctgaaag aacagaaaca aaactatcag cagaatattg agatgtttta ctaagttgta   16320
tatctatact gcttgtaatt tttaccccaa gcaagaatta cttttttggaa aaagaaaatt   16380
caggaaataa agcatttctt taaacttcat gtttaaacaa atggtgatgg aataaaaagag   16440
ttcttattca tcataaacac acacagcaca catgcacgca tgtgcgtgaa cacaccctt    16500
acttgataaa taccatgttg aatattttag tctttcctt taggttctat cccttcactc    16560
aaaatgcggt tataaataaa tgtactttc atgtgccttc tgcctaaacc cactttaata   16620
taacttaca gtcccattat cattatatgtc tcaaagctag actcagcctg aaactacccct  16680
```

```
ttcatttgga acccttatta aaatgccaca tacagctcct tcaaataaaa acaaacccta 16740
ggacctgaca ctaggcttcc tttgttgcta ctcataatgg ccaagttctg tgcttataat 16800
acatcttctt tcattttatt gctacatatc caagggtttt atatgttttt cttattatat 16860
cttaattcaa aacaccatca cgctcttttc cagatgaaaa taaggaaaag aaattgagca 16920
actgactgac ttaaaggtca taaaactata tagtagcaga gtcagcaaaa gaagaaacac 16980
acatctccca agtagaggct gaaaaccagt accattcacc tccagggtga gctatataca 17040
gattacaaag tcaccttctc taaatgttca aactgaatcc cataccata ctttaccact 17100
acctcgtaag aacagcctca gatcttgtta tagccttttt tttagcatgc tgaagccaat 17160
aaaatgcttc ccattcagca agagaaacaa gttctgaaac actgaataat ctgcccaggg 17220
cctatgaaca tttccactgt agacattcaa ttggattcta tgatgcacag cttaggagtc 17280
ctccacacag gcagaacatt agataaatc ttggattcta tgatgcacag cttaggagtc 17340
tgtttagcac aatttaagtc caaatagtta ttaaatcctc ctctgttcca gaaacagtgc 17400
taaatactgt gaatataaaa attgaaaaga tactctcctg gctcccaaga aagtcagcca 17460
gatagaggag acacaggcac acaaatcact gtcacatgaa gctctacctc cctaacttca 17520
ttcaatactt tatcttccct tagaaaaactc ttctcccttg gaaatttatt tgcatttcta 17580
aataccattc cttactaaaa ggaagcaggg ctccttgggg aaatagctga ttctaggtgt 17640
ggactatgaa atgaaaatgg tgagtctggg acatcccatg ttgcccagaa atcaaggaac 17700
tgcccaaaga ttaacagagt catgttaaat ggacctaaga gtgaaccaga aggagctcac 17760
tttgccccgc gtggaacaat ttcaagaaaa acatgacagt aatgaattat aaaacatgaa 17820
ttaaaataca tattggtact aaaaagagaa caaaaggatg tggctttgga taaagctctt 17880
cttcatggag gaataccagc taataaatgt aaaggaaatg agagaattag aaaaattatc 17940
attttgtaaa ccttaatata ttcacctaga catgctaaaa ccactgagta aaaggctgct 18000
tgggaagagg atgctcacat gatctcagag tttcacacca cagataattt attagataca 18060
ggaaggagga tgtgatcaag cttctgtgga cccccagcca ggccccacaa cactatgtgc 18120
ctccttgtga tgtgggagct acacagcatc gcccacacag cttctcgcca aaactgtttg 18180
aagctaatca caagggaaga actggacagc ttctgaccat gagacgctcc accagacaac 18240
ttgcttggcc tctccaaaga aacttgcttg gcctctccaa agaaaactca gtttcattta 18300
aaaacaaaac taattattta aaaacaaacg aaaagcaagt tgtggacttg agctccaggg 18360
acagagcaga catactttc cctgttcttc ccagtaagtg gtaataaaaa ccctcaacac 18420
tagatataaa acaaatataa gaaggttctg gaagggaag aggaggcaga ctatccaggt 18480
gccttgaggc ccacagaaca acccagtgat gggttcactg ggtcttcttt ttgcttcatt 18540
atctcagact tggagctgaa gcagcaggca acttcaaaac accaagggc acagattgaa 18600
aagcccaag aaaagcctgc cctctctagc caaaggacca ggaaggagac agtctaatga 18660
gatggaacac atttagacag taactgccca tttaccagca ataactgagc aggagctca 18720
gacttccagt cttgtgagga cgtaccaagg tacccaacac ccccaccaag gctgagtaag 18780
gactgcgact tttatccctg catggcagta gtaaggagcc catcctcac tggcagccac 18840
tgtcagggga acctggactt ccactcccac ccaggagtga tgaggccctc cctgctgggg 18900
tcatgtcaga ggaggcctag tggagattca gtgacttaac cttttcccag agataatgag 18960
gccacctttc ctccctcttc cccatggtg acagtgaaag cactgtggca agcagtaggc 19020
actcctaccc ctcctagcca gggaggtatc agggaggcca agtagggaac cagaataccc 19080
acaaccaccc agcagcaaca gggtccccc acccattgg gtgtcaatgg aagcagagcg 19140
gaaagcctgg atatttaccc ccatctagaa gtaacaagct gatgtccccc ttcttctact 19200
acaatggtgt tcaaaacagg tttaaataag gtctagagtc tgataacgtca atacccaaat 19260
cgttgaagtt ttcattggat atcatttata ccaagagtca ggaagatccc aaactgaaag 19320
agagaaaaga caattgacag acactagcac taagagagca cagatattag aactacctga 19380
aaggatgtta aagcacatat cataagaactc aacagctgg gcgcggtggc tcacgcctgt 19440
aacccagca ctttgggagg ccgagacagg tggatcacaa gatcaggaga tcgagaccat 19500
ggtggtgggc acctgtagtc ccagctctac gggagcctga ggcaggagaa ttgcttgaac 19560
ctgggaggcg gaggttgcag tgagccgaga tcgcgccatt gcactccagc ctgggcgaca 19620
gagcgagatt ccgtctcaaa aaaaaaaaat agcaaggcat ggtggtgggc acctgtagtc 19680
ccagctctac gggaggctga ggcaggagaa ttgcttgaac ccgggaggcg gaggttgcag 19740
tgagccgaga tcgcgccatt gcactccagc ctgggcgaca gagcgagatt ccgtctcaaa 19800
aaaaaaaaat agcaaggcat ggtgcttgaa caaatgaaaa aaaatcttg gcaaagaaat 19860
aaaagatata tatttggcc agtgcagtg gctcacgcct gtaatcccct gcactttggg 19920
aggctgaggc aggcggatca cctgagtca ggagtttgag accagcctga ccaacatgga 19980
gaaaccccgt ctctactaaa aatacaaaat tagccagtca tggtggcaca tgctgtaat 20040
cctagctact caggaggctg aggcaggaga atcgcttgaa cccaggagg gaggttgcgg 20100
tgcacattgc actccagcct gggcaacaag agcaaaactc catctcaaaa aaaatagac 20160
atatttaat ggaaattta gaattgaaaa atacagtaac caaattgaat ggaaagacaa 20220
catagaatgg aggggcaga caaataatc agtgaacttc aacagaaaat aatagaaatt 20280
acccaatatg aagaacagaa agaaaataga ctggccaaaa aataaagaag aaaaaagagg 20340
agcagcagga ggaatgatgg aaaaagagaa aggaagagag gaagggaagg aggggaggca 20400
ggagtgaggg agaaagtctc aaagacctct gagactaaaa taaaagatct aacacttgtc 20460
atcagggtcc aggaaagaca caagatggc acagctgaaa acgtattcaa aaataataag 20520
ctgaaaactt cccaaatttg gcaagagaca taaacctata gattcgaaat gctgaaccc 20580
aaataaaaag cccaataaaa tccaccacaa aatacatcat agtcaaactt ctgaaaagac 20640
gaaaagagaa aacgtcttga aagcagtgag tgaacaaaca cttcatgtat aagggaaaaa 20700
```

19

```
caattcaagt aacagatttc ttacagaaat taaggaagcc agaaggaaat gacacaatgg    20760
ttttcaagtg ctgaaagaaa agaagtgtca acacaaaatt ctagattcag taaaaatatc    20820
cttcaagaat caatgggaaa tcaagacagt ctcagataaa gcaaaataag agaatatgtt    20880
gccagcagat ctcccctaaa ggaatggcaa aaggaagatc atgcaacaga ccaaaaaatg    20940
atgaaagaag gaatccagaa acatcaagaa gaaagaaata acatagtaag caaaaataca    21000
tgtaattaca ataaaatttc tatctcctct taagacttct aaattatatt gatggttgaa    21060
gcaaaaatta taaccctgtc tgaagtgctt ctactaaatg tatgcagaga attataaatg    21120
gggaaagtat aggtttctat acctcattga gatggaagaa tgacaacact gtgaaaagtt    21180
acatacacac acacacgtaa gtatatataa atatatgtgt gtatatgtgt gtgtatatat    21240
atatatacat ataatgtaat acagcaacca tataccaaga gataataacc    21300
aaaaacaatt tagataaatt gaaatggaat tctaaaaaat attcaaatac tctacaggaa    21360
tcaattaaaa gacagagata aaagagaagg acaaactaaa tttttaaaaa acataaataa    21420
aatggtagac ttaagcccta acttatcaat aattcataa atgtaaatga tctaattata    21480
tcaattaaaa gacagagata gcaagggtaa tcccagcaact tcgggaggcc aaggcgggtg    21540
tgggctgagt gcagtgactc acacttgtaa tcccagcact tcgggaggcc aaggcgggtg    21600
gatcacctga ggtcaggagt tccagaccag cctggccaac atggtaatac cccatctcta    21660
ggaggctgcg acacaagaac tgcttgaacc cgggcagcag aggtagcagt gggccaagat    21720
tgcgccactc cagcctgacc gacagagtga gactccacct cagttgaaaa acagaaaaga    21780
aacctgcttt aaatatacca acatatgttg gttgaaatta aaagaataaa atatatcatg    21840
aaaacattaa tcaaaagaaa gggagtggctc tattaataac ataaaaataga cttcagagaa    21900
aagaaaattt caagagacag gaataaaagg atcaagaaaa gatcctgaaa gaaaagcagg    21960
caaatcaatt attctgcttg gagattcaac accctctctt aacaactgat agaacaacta    22020
gacaaaaaaa tcagcatgga gttgagaaga acttaacacc actgaacaac aggatctaat    22080
agacatttac ggaacactct acccaacaat agcaaaataa acatttcttt caagtattca    22140
ctgaacatat ccttagaccc taccctgggc cataaaacaa agctcactag tgattgccga    22200
aggcttggat ggacagtgga agagctgcat ggggagggag aaggtgacag ttaaagagtg    22260
taggatttct ttttgggata atgaaaatgt tccaaaattg attgtggtga tgttggcgca    22320
actctacaaa tataaaaaag gccattgaat tgtacgtttt aagtgggtga aacatatggt    22380
atgtggatta tatctaacgc tttttaaaaa cttaacacat ttcaaagaat agaagtcata    22440
cagagtgtgc tctactggaa tcaaactaga aagaggtaac tggaggataa cgagaaaagc    22500
ctccaaatac ttgaaaactg gacagcacat ttctaaaatc atccgtgggt caaagatatt    22560
catttctgat attcattttt attgtttaat gtattttttaa aaatttctta aagaatacaa    22620
actgactaaa aatgaatatg gctgggtgcg gtggctcacg cctgtgatcc cagcactttg    22680
ggaggccgag gctggtggat cacaagatca ggagttcgag accagcctgg ccaagatggt    22740
gaaaccccgt ctcaactaaa aaactacaaa aagtagccaa gcgcagtggc gggagcctgt    22800
ggtcccagct acttgggagg ctgaggtagg agaatcgctt gaacacaggc agcagaggtt    22860
gcagtgagcc aagattgtgc cactgcacgc cagcctgggc gacagagact gcctcaaaaa    22920
aaaaaaaaaa aaaaagaata tcaaaatttg tgggacataa ttaaagcaat gctgagaggg    22980
aaatttataa cactaaatgt ttacattaga aaagagaaaa agtttcaaat caatagtctc    23040
cactcccatc tcaagaacac agaagatgaa gagcaaaata aacccaagc aacgacaaaga    23100
aagaaaatat aaaaataaat cagtaaaatt gaaaacagaa acacaataaa gaaaatcagt    23160
gaaacaaagt actgattctt cgaaagatta ataaaattga caaacctcta gcaaggctaa    23220
caaacaaaaa agaaagagaa cacggattac cagttattag aatgaaagca taattagaaa    23280
caactctaca cattataaat ttgacaatgt agataattgg gatgcctga tagactcatga    23340
caattacca caactcaccc aatatgaaat agataattgg gatgcctga tatatttttat    23400
gaaaattgaa tttgtaattt taacactctt aaacacgaaa cattaaactt aatattttat    23460
aaatattaga taaggtaatt atacctctcc ttaacaaata aaaacgacaa attattttac    23520
agctaaagay atgtatgtac tgtgaaaaat atcttcagaa aaatagaact ttgtttgaag    23580
aataaggatt taaaaaatgt ttttaactct caagaagcaa atatctgggc ccagtggtt    23640
tcactgaaga attctaccaa atgtttaatg aagaattacc accaactcta catgcatct    23700
ttgagaaaac tgaagagaag ggaacatctc ccagttcatt ttatgaagtg ggtgtttactc    23760
tgatactaga actgtataga gacagctact cttgacacac tgcctatggg tagctctgct    23820
ctgcaggacag agtcagaaaa aaaaaaaaaa gaagcactgg acaagggcag tataaaaaa    23880
agaaactggg ccaggtgcag tggctcacac ctgtaatccc agcactttgg gaggctgacg    23940
ctggtggatc acctgaggtc aggagttga gactagcctg gccaacatgg taaaaccctg    24000
tctctactaa aatacaaaaa ttagccaggc agggtggtgg ggaaaataaa aaggaaaaaa    24060
aaacaaaaat aaactgcaga ccaatatcct tcatgagtat agacacaaaa ctccttaaac    24120
tccttaacaa aatattagca agtagaagca atatataaaa ataattatac accatgatca    24180
agtgggactt attccagaaa cgcaagtctg gttcaacatt tgaaaacaag gtaacccact    24240
atatgaacgt actaaagagg aaaactacat aatcacatca atcaatgcag aaaaaagcat    24300
ttgccaaaat ccaatatcca ttcatgataac tctaataaga aaataagaa tacagcaatt    24360
attccttgac ttgataaagc ttacaaaaga ctacaaaagc ttacagctaa cctatactta    24420
atgtgaaaa actaaatgct ttcccctacg atcaggaaca aagcaaggat gttcactctc    24480
attgctctta tttaacatag ccctgacgtt ctaacttgtg caaaacgata agaaagggaa    24540
atgaaagacc tgcagattgg caaagaagaa ataaaactgt tcctgtttgc agatgacatg    24600
attgtctcat agaaaatgta aagcaactag gggtaggggg gcagtggaga cacgctggtc    24660
```

                                    20

```
aaaggatacc aaatttcagt taggaggagt aagttcaaga tacctattgc acaacatggt    24780
aactatactt aatatattgt attcttgaaa atactaaaag agtgggtgtt aagcgttctc    24840
accacaaaaa tgataactat gtgaagtaat gcatacgtta attagcacaa cgtatattac    24900
tccaaaacat catgttgtac atgataaata cacacaattt tatctgtcag tttaaaaaca    24960
catgattttg gccagcgcaca gtggctcata cctgtaatcc cagcattta ggaggctgag    25020
gcgagcagaa aacttgaggt cgggagtttg agaccagaat ggtcaacata gtgaaaccct    25080
gtctccacta ataatacaaa aattagcagg atgtggtggc gtgcacctgt agacccagct    25140
acttgggagg ctgaggcacg agaattgctt gaacaaggga ggcagaggtt gcagtgagct    25200
gggtgccact gcattccagc ctggtgacag agtgagactc catctcaaaa aaataaaat    25260
aaagcatgac ttttcttaaa tgcaaagcag caggcagatc acaaggtcag ggttcgaga   25320
accactttgg gaggccgagg caggcagatc acaaggtcag ggcttgaga ccagcctgac    25380
caacatggtg aaaccccatc tctactaaaa aatatataaa ttagccaggc atgtgtagtc    25440
tcagctactc aggaggctga ggcaggagaa tcacttgaac ccgggagcag aggttgcagt    25500
gttgagccac cgcactccag cctgggtgag agaacgagac tccgtctcaa aaaaaaaaag    25560
caaaataacc taattttaaa aacactaaaa ctactaagtg aattcagtaa gtctttagga    25620
ttcaggatat atgatgaaca tacaaaaatc aattgagctg gacaaaggag gattgtttta    25680
ggtcagtagt ttgaggctgt aatgcacaat gattgtgcct gtgaatagct gctgtgctcc    25740
agcctgagca gcataatgag accacatctc tatttaaaaa aaaaaaaatt gtatctctat    25800
gtactagcaa taagcacatg ggtactaaaa ttaaaaacat aataaatact gttttttaatt    25860
gcctgaaaaa aatgaaatac ttacatataa atctaacaaa atgtgcagga cttgtgtgct    25920
gaaaactaca aaacgctgat aaaagaaatc aaagaagact taaatagcgt gagatgactaaa    25980
atgcttatag gttggaaaac ttaatatagt aaagatgcca attttatcca aattattaca    26040
atacaaaaat gtatacggaa atatgcaaag gaactagagt agctaaaaca aatttgaaaa    26100
agaaaaataa agtgggaaga atcagtctat ccagtttcaa gacttacata gctacagtaa    26160
tcaagactgt gatattgaca gagggacagc tatagatcaa tgcaaccaaa tagagaacta    26220
agaaagaagc acacacaaat atgcccaaat gatttctgac aaaggtgtta aaacaacttca    26280
acgggggag atatgtctct cattaaaggg tgtagagtca ttgcacatct ataggcaaaa    26340
agatgaacct gaacctcaca ccctacagaa aaattaactc aaaatgactc aaggactaaa    26400
cataagatat acatctataa aacatttaga aaaaggccac gcacggtggc tcacgctcgt    26460
aatcccagca ctttgggagg ccaaggcagg tggatcacct aaggtcagga gtttgagacc    26520
agcctgatca acatggagaa gccccatctc tactaaaaat acaaaattag ctggacgtgg    26580
tggcacatgc ctgtaatccc agctacttgg gaggctgagg catgagaatc gcttgaacct    26640
gggggggcaga ggttgcggtg agccaagatc acaccattgc actccagcct gggcaacaag    26700
agcaaaactc caactcaaaa aaaaaaaaaa aaaggaaaaa tagaaaatct ttgggatgta    26760
aggcgaggta aagaattctt acacttgatg ccaaactaag atctataagg ccagtcgtgg    26820
tggctcactgc ctgtaattcc agcactttg tcaactagat gaaaggtata tggaagtcac    26880
ctgtattatt ctttcaactt ttctgtaggt ttgacatttt tttagtaaaa aattggggga    27000
aagacctgac gcagtggctc acacctgtaa tcccagcact ttggggagcc ggggcaggtg    27060
gatcacacgg tcaggagttc gagaccagcc tggccaacat ggtgaaaccc cgtctctacc    27120
aaaaatataa aaaattagcc gggtgtcatg gtgcatgcct gtaatcccag ctactgagga    27180
ggctgaggca ggagaatcac ttgaacctgg gaggtggaag ttgcagtgag ccgagattgt    27240
gccactgcac tccagccttg ggtgacagga cgagactccg tctcaaaaga aaaaaaaaaa    27300
aaagaaatatc aaacgcttac tttagaaact atttaaagga gccagaattt aattgtatta    27360
gtatttagag caatttttat gctccatggc attgttaaat agagcagccc tatcaatatc    27420
agtggagttc aacagctgtt aaatttgcta actgtttagg aagagaggcc catcatatcc    27480
actgtcattt gaggctgaca ataagcacac ccaagctgtc acctccttga ggagcaacat    27540
aagggggttta accctgttag ggtgttaatg gttggatat ggtttgtttg gccccaccga    27600
gtctcatgtt gaaattttgt ccccagtact ggaggtgggg ccttattgga aggtgtctga    27660
gtcatgtgggg tggcatatcc ctcctgaatg gtttggtggcc attcttgcag gaatgagtga    27720
gttcttactc ttagttccca caacaactgt ttattaaaaa cagcctggca ctttccccca    27780
tctctcgctt cctctctcac catgtgatct cactggtcc ccttcccttt atgcaatgag    27840
tggaagcagc ctgaagccct cgccagaagc agatagtgat gccatgcttc ttgtacagcc    27900
tacaaaacca tgagcccaat aaacctttt tcttttataa ttatccagcc tcaggtattc    27960
ctttatagca agacaaatga accaagacag ggggaaatca acttcattaa aataatctat    28020
gcagtcacta aacaaataag aacaagaggc tccagaagtg ggaagccaat acccagagtt    28080
cctacaatac agtatctgaa aagtccagtt tccaaccaaa aaatatatat atacaggccg    28140
gacatggtag cttatgtctg taatcccagc actttggagt gctgaggcgg gcagatcacc    28200
ctaggtcagg agttcgagac cagcctggcc aatatggcaa aaccccgtct ctactaaaaa    28260
tacaaaaatt agccaggcat ggtggtggat gcctgtaatc ccagctactc gggaggctga    28320
ggcagggaat cacttgaacc caggaggcag aggttgcagt gagccgagat cacgccactg    28380
aactccagcc tgggcaacaa agtgagactc cacctcaaaa aaaaaaaaaa tatacatata    28440
tatatgtgtg tgtgtgtgtg tgcgcgcgtg tgtgtatata cacatacaca tatatacata    28500
tatacagaca cacatatata tatgaagcat gaaaagaaac aaggaagtat gaaccatact    28560
ttctgtggtt atgatgagtt gggggtatcac gggggaagta gacaagggaa actgcaagtg    28620
agagcaaaca gttatcagat ttaacagaaa aagactttgg agtaaccatt ataaatatgt    28680
ccacagaatt aaagaaaagc gtgattaaaa aaggaaaagga aagtatcata acaatattac    28740
```

21

```
tccaaataga gaatatcaat aaaggcatag aaattataaa atataataca atggaaattc    28800
cggagttgaa aggtagaata actaaaattt aaaattcact agagaaggtt caacactata    28860
tttgaactgg cagaagaaaa atttagtgag acaaatatac ttcaatagac attattcaaa    28920
tgaaaaataa aaagaaaaaa gaatgaagaa aaataaacag aatctcagca aacattaaag    28980
caccattaat cacattaaca tatgcatact gagagtaccg gaagcagatg agaaagagga    29040
agaaaaaata ttcaaatgat ggccagtaac ttcctagatt tttgttttaa agcaataacc    29100
tatacaatca agaaactcaa tgaattccaa gtaggataaa tacaaaaaga accacaaaca    29160
gatacaccat ggtaaaaatg ctgtaagtca aacaaaaagaa aaatattgaa agacgctaga    29220
ggaaaactta taagagaacc tcacttacaa aagaacatca cttataaaag aaccacaata    29280
atagaaacag ttgacctctc atcagaaaca atgaatgata acatatttga agtgctcaaa    29340
gaaaaaaaat aaagattcct atatacgaca aagctgtctt tcaaaaatat acatccaaaa    29400
ggattgaaac caggtgtctg aagagttatt tgtacatcca tgttcatagc agcattattc    29460
acaatagcca aaaggtagaa gcaacccaag ggtccatcga caaataaata aaatgttggta    29520
tatgtataca caatggaatt tattcagtat taaaaaggaa tgaaattctg acacatgcta    29580
caacatggct aaaccttgag aacactatgc taagtgaaat aagccagcca caaaaggaca    29640
aataccatat tacttcactt gtatgaaata cctagggtag tcaaattcag agatagaaag    29700
taaaacaagtg gttgccaagg gctgagggag ggagtaacgt ggagttattg ttgaatgggt    29760
acagaatttc agttttgcaa gataaaaaga gttctggaga cagatgtgg tgaggggtggt    29820
acaacaatac aaatatactt tatactactg aacagtatac ttaaaaatga ttaaaatggt    29880
gaaaccccgt ctctactaaa aatacaaaaa aattagctgg gtgtggtggc gggcacctgt    29940
aatccagact acttgggagg ctgaggcagc agaattgctt gaaaccagaa ggcgggaggtt    30000
gcagtgagct gagattgcgc caccgcactc tagcctgggc aataagaca aaactccgtc    30060
tcaaaaaata aaaaataaaa aaaatttaaa aatgattaag caggaggcca gacacggtgg    30120
ctcacaccta taatgccagc actttgggag gccgaggcag gcgatcactt gagaccagga    30180
gtttgaacgc agcctggcca acatggcaaa accctgtctc tgctaaaaat acaaaaatta    30240
gccaggcatg gtggcatata cttataatcc cagctactgg tgagactgag accagagaat    30300
tgcttgaacc caggaggcag agattgcagt gagtcgagat cgcgccactg aattccagcc    30360
tgggcgacag agcaagattc tgtctcgaaa aaacaaaaac aaaaacaaaa agcaaaacca    30420
aaaaataatt aagcagaaaa cgagattgct gctgaggaag agaaagatgt gcaggaccaa    30480
ggctcatgag agcacaaaac ttttcaaaaa atgtttaatg attaaaatgg taaatttat    30540
atgtatctta ccacaaaaaa aaggctgggg gggcagaaa tgaaggtgaa ataaagacat    30600
cccagagaaa caaaagtaga gaatttgttg ccttagaaga aacaccacag gaagttcttc    30660
aggctgaaaa caagtgaccc cagagggtaa tctgaattct cacagaaaat tgaagcatag    30720
cagtaaaggt tattctgtaa ctatgacact aacaatgcat atttttcct ttcttctctg    30780
aaatgattta aaaagcaatt gcataaaata ttatatataa agcctattgt tgaacctata    30840
acatatatag aaatatactt gtaatatatt tgcaaataac tgcacaaaag agagttggaa    30900
caaagctgtt actaggctaa agaaattact acagatagta aagtaatata aggaaggaact    30960
taaaaataaa attttaaaaa atttaaaaat aataattaca acaataatat ggttgggttt    31020
gtaattataa tagacataat acaaaaatac cacaaaaagg gaagaagaca atagaactac    31080
ataggaataa cattttggta tctaactaga attaaattat aaatatgaag tatattctgg    31140
taagttaaga cacacatgtt aaaccctaga tactaaaaag taactcacat aacatacaggg    31200
aaaaataaa taaaataatt aaaatgtttg tattagtttc ctcagggtac agtaacaaac    31260
taccacaaat tgagtggctt aacacaactt aaatgtattt tctcccagtt ctggaggcta    31320
aacacctgca atcaaggtga gtacagggcc atgctccctg tgaaggctct agggaaagaat    31380
cctccttgt ctcttccagc ttccagtggt tctcagtaac cctaagtgct ccttggcttg    31440
tagctatatc attcctagca accagaaaga agaaaataat aaagattatg gcaaaaaata    31500
atgaaatcaa aaggagaaaa atggaaaaaa ataaataaaa ccaaaagcta gttctttgaa    31560
aagatcaacc aagttaacaa accttttaac tagactgaca aaaaggaggt aagactcaaa    31620
ttactagaat cagaaataaa agaggggaca ttactaatga gggattagaa aagaatacta    31680
cgaacaaatg tgtgccaaca aattagaaaa cttagtagaa atggacaggt tcctaggaca    31740
acatcaacta ccaaaattta ctcaagaaga aagagacaat ttgaatgagc tataacaagg    31800
gaagagactg aattgacaac caagaaacta tccacaaaga aatcccagg cccagaagat    31860
ttcactgtga aattctttca aacttataaa tataaattaa catcagttct tcacaaaactc    31920
ctccaaaaaa aagaacagat ctctatttac aggcgataacg atctttagaa aatcctaagg    31980
gaactactaa gacactatga taactgataa acaagttcag caaggctgca ggatagaaaa    32040
ccaatataca aaaatctatt atatttctat acacttgcag tgaacaaccc aaaaatgaga    32100
ttaagaaaat aattcaattt acaataacat caaaaagaat aaaaacactc aaaaataaat    32160
ttattcaagt aagtgcaaaa cttatactct agaagctaca aaacactgtt aaaagaaaatt    32220
aaaggttac ataaatgaaa aactatccca tgttcatgga tcaaaagact tattactggc    32280
aatgctctcc aaattgatct ataaattcaa caaaatcctt atcaaaatcc cagatgaggc    32340
tggggggtgc ggttcatgcc tgtaatccca gcactttggg aggctgaggc acgcagatta    32400
cctgaggtcg ggagctcgag atcagcctga ccaacatgga gaaaccctat ctcttccaaa    32460
aatacaaaat tagtcaggcg tggtggcaca tgcctataat cccagctact cgggaagctg    32520
aggcaggaga atcgcttgaa cccaggaggc agaggttgca gtgagccaag atcgtgccaa    32580
tgcactccag cctgggcaac aagacaaaa ttccatctca aaaaaaaaa aaaaaaaatc    32640
ccagatgact tcactgttga aattgaaaag attattctaa aattcacatg gaattgcaag    32700
accttgagaa tagccaaaac aaacttgaaa aacacgaaca aaatatagga tgactcactt    32760
```

                                                    22

```
gccaattgca aatgttacga cacagcaaca gtaatcaaga ctgtgtggta ctggcaaaag    32820
acacatacat acatacatat caatggaata taattgagag tacagaaaca agcctaaaca    32880
tctatggtaa gtgcttttct attttttttct tttttttttt cttttttgta gagatagaat  32940
ctcaccatgt tgcccagcct ggtcttcaac ttctgggctc aagcaatcct cccactgtgg   33000
cctcccaaag tgctgggata actggcatga gccaccacat ccagcccaga tgatttttcaa 33060
aaaagtcaac aagaccattc ttttcaacaa ataggtctgg gatgaccaga tagtcacatg   33120
aaaaaaaaaa tgaagttgga ccctccatca cactaaaagtg ctgcgattat aggcatcagc 33180
caccacatcc agcccaaatg attttcaaaa aggtcaacaa gaccattctt ttcaacaaat   33240
aggtctggga taatcagata gtcacatgaa aaaaaaaatg aagttggacc ctccatcaca   33300
ccatatgcaa aaattaattc aaaaatgaat tgatgactta aacgtaagag ttacgactgt   33360
aaaactctta gaaggaaaca tacgggtaaa tcttaaagac gttaggtttg acaaagaatt   33420
cttagacatg acaccaaaag catgaccaac taaggtaaaa tagggtaaat tgtacctacc   33480
aaaatgaaaa acctttgtgc tggaaaggac accatcaaga aatggaaagc caaaatagcc   33540
aaggcaatat taagcaaaaa gaacaaagct ggaggcatca tactacctga cttcaaagca   33600
acagtaacca aaacagcatg gtactagtag aaaaacagac acatagacca atggaacaga   33660
ataaagaacc caaaaataaa tccacatatt tatagtcaac tgattttttga caatgacacc   33720
cctcaataaa atgatactag gaaaactgga tatcgatatg cagaagaata aaactagacc  33780
cctatctctc accatataga aaaatcaact cagactgaat taaagacttg aatgtaagac   33840
ccaaaactat aaaactactg gtagaaaaca taaggaaaaa cgcttcagga cattggtccca 33900
ggcaaagatc ttatggctaa aacctcaaaa acacaggcaa caaaaacaaa aatggaaaaa   33960
tagcacttta ttaaactaaa aagctcctgc acagcaaagg aaacaacaga atgaaaagac   34020
aacctgtaga atgggagaaa atatttgcaa actatccatc catcaaggga ctagtatcca   34080
gaacacacaa gtgactaaaa caactcaaca gcaaaaaagc aaataatctg gtttttatat    34140
gggcaaaaga tctgaataaa cattctcaaa ggaagacata caaatgtcac tatcattctg   34200
ccagtaccac actgtcttga ttactttgtta gtgtataaat ttttaaattg ggaagtgtga  34260
gtcatcctac actttgttct tgttttttcaa gtttgttttg gctattctgg gagccttgca  34320
agtataaaat agccaacaag tatgaaaaaa tgctcaccat cactaatcat cagagaaata   34380
aaaatcaaga ccactagag atatcctctc actccagtta gaatggctac tatcaaaaag   34440
acaaaatata atggatgctg gcaaagattt ggagaaaggg gaactcctat acactgtggg   34500
tagggatgca aattggtaat ggccattatg gaaaataata ctgaggtttt tcaaaaaact   34560
gaaaatagaa ctaccatatg atccagcaac cctactactg ggtatttatc cagaagaata  34620
aagtcagtat actgaagaaa tatatgcact ctcatgttaa ttgcaacact gttcacaaca   34680
gccaagacag ggaataaatc taaatgtgca tcaacagatg aatggataaa gaaaatgtgg   34740
catatacact caatagaata ctattcagcc attaaagaag aatgaaatcc tgtcatccca   34800
gcaactgga tgaacctgga ggacattata tttaatgaaa taagtaaagc acaaaaagat   34860
aaacagtaca tgttctcact cagacatggg tgctaaaaag aaaatggggt cacagaatta   34920
gaagggagg cttgggaaaa gttaatggat aaaaatttac agctatgtaa gaaaataaaa  34980
ttttagtgtt ctatagaact gtagggcgag tatagttacc aataacttat tgtacatgtt   35040
caaaaagcta gaagagattt tggatgttcc cagcacaaag gaatgataaa tgtttgtgat   35100
gatggatatc ctaattaccc tgattcaatc attacacatt gcatacatgt atcaaattat   35160
cactctgtac ctcataaaata tgtataatta ttacgtcaac aaaaaaagga aaaaaaagaa  35220
aattaagaca acccacataa tggaagaaat aaaaatatctg caaattatat atatctgata  35280
aatatttaat attttataata tataaagaac tcctacaact caagaacaac aacaaaacaa  35340
cccaattcaa aaatgggtaa aagccttgaa tatacactta tctaaagact atataatcaat   35400
ggccaataaa gacacgaaaa gatgctcaac atcactagtc atcagggaaa tataaatcaa   35460
aaccacaatg tagaatgtag acaccacttc atatgcacta ggatggctag aataaaaaggg 35520
taataacaaa tgttggtaag gatgtgaaaa aatcagaaac ctcattcgct gctgttgggga 35580
atgtaaagtg atgcagccac tttggaaaac agtctggcag ctcctcaaat tattaaaatac 35640
agagttaccg tatgacccag gaatattcct cctgggtcta taaccaaaaa aatgaaaaaca 35700
tatatccaca taaaaactg tacatggcca ttttatagcaa cattattcat aacagcaaag  35760
gtggtaagaa cccatatgcc catcatctga tgaacaggta aataacatgc ggtattatcc   35820
atacactaga atattatctg cccatacaag gagtgacatc cagctacatg ctacaaggat   35880
gaatctcgga aaccttatgc taagtgaaag aagccagtca caatgacca cagattatga   35940
ttccatgcat cggaaatgac cagaatagggg aaatctatag agacagaaag tagattagtg  36000
gttgggtggg gctgggagga caggtagtac actactttttc cagaactact ggaacaaagt  36060
accacaaact ggggagctta aacatagaaa ttgatttcct cacagttctg gagactaggga 36120
ctctgagatc aaggtgtcag cagagctggt tcttttctgag ggccctgagg caaggctctg  36180
tcccaggcct ctctccttgg ctggcaggtg gccatcttct ccctgcgtct tcacatcatc   36240
ttttctctgt gtgtgcccat gtccaaatt tgattggctc attctgggtc atggccaatt   36300
gctatgcaca aagtgaagtc tacttccaaa agaagggaag agggaacact gacctgtagg   36360
aacttatagt cattttaatg tccgcttttc ctatgagatt gtgaacacac agaagtaggg   36420
ttttttatcta cattgtgcaa agtttaataa gaaaaataga attcaagaga agcagttcaa  36480
tagcaggaat ttaatatgga aactaattac aaggtttagg gcaggactaa aaagccagtt   36540
gggatgggta gccaacccag agattagcaa cagtgggacc cctcattacct accacccatg 36600
aagctggaag gataaaggag gggctattat cagagtccac aagccagtgt cagagtccttt  36660
ggctggagct gggaccaccc tagagacact gtgcaaagca gaaaacaagg gggaaaaacc   36720
ctgacttctc ccttcctccc acctttcaat ctcccactag tgcttcctac tagccatact   36780
```

23

```
tggccagaga cagtgacaag gaacactgca aaatgaagtt tgtaggaatc atctccctct    36840
gagacgagaa aatatggaag ggtagaaaat gaatcagagg ataaagagaa aaaaccctga    36900
gtactatctt atttatcttt gtatctccag tgcctaatct gtctctcaaa aaaggaaagc    36960
aattgagaga aactgaaaac tccaattgaa atgaaagaat ggagaattac tggactagaa    37020
gagaagagaa aaatttattc cgcatagagt aaacaagaat ggattcacaa aggacgtgat    37080
gaatgaaaag ctataatcag caaagatttg ccagagaaat taaaaagtgg taaactcagc    37140
cacgctgtac aacctgaagg cacaatgcat gaaaacgttt caagaaatga caagatttga    37200
agtcaaattc taagtgcttt tccagaatct ctcaaagcaa tctataagct acccccatttt    37260
attaaataaa atggaaactt actaaacttt cccccttgtat taaactaaca tatgtcctaa    37320
tagcaaacga ttctggaatt cctagagtaa aatataattc gtcaaagtgt attgctcttt    37380
taatattctg tgacctcct tttgctattt aggatatttg tatacacatc acacgtaaat    37440
ttggtctata gtttacatct acgggcttat actgttcttt ttttcatttt tttaaaattt    37500
ccaacccca gtatccatat actgctctct atcagggtta ttttaacttt gtaaaatcag    37560
ctgagatgct ttccatgttt tttttttta ttttctgcca catttgaata gcataggagt    37620
taccaccatc aacctttgat tatttaagca ttcacgattc cacgtgtgga tttttattc    37680
agagtctttc ttgtcattcc tgctatcagc acagaaccca atctcagctt tccagctata    37740
ctctcacccc ccttcattta gacatcacct tcttctagaa cgtcttacct gacatgccct    37800
gctcccaacc cctgctgccc aattgtgtgc tctcccgtgt cctggcctgc catcctcttt    37860
agtaattgcc tgctcctca tctgtctccc cacccagaca ttaagctgaa tagactggat    37920
ttgtgtcttg tccatcacta taatctcagc acctagtacc tagtaggtac ttaccatgta    37980
ttcattagca aaatgttatg tataacttg caccttaaaa acaagagaag gaagacaaaa    38040
ttaagtctta agactatggt ttagaacatg gatcagaaac tacagtctgc agcccaaatc    38100
cagaccaaat gaagagacca tgttcatttta cacacaacct atagcagctt tcacactaca    38160
ggagcgagc taagtagttc caagggaaca cacggccctg caaagcctaa aatatttact    38220
ctatagctct tcacagaaaa agttttcaga tccctcgttt agaactcttg ttcatatgca    38280
atttcactaa accatagttt tttgggtttg tttgttttt ttttggcaaa aggaatgagc    38340
cgatccagaa aaggttgaaa agaatgaatc attactgctg aaagaaatgtg cacacagtcc    38400
gtcagtattc tgctgccatg ctgacaccca tccaatagtg tcatgagatg cagcagctac    38460
tactgtgttc tcaatgccga gtccacccac tccataacca tgtccaagca atcttgggaa    38520
catcatcacc atgcttgttt atccttaagg tattgcctca catacagcag tggctggtca    38580
taaagtcaaa tgacactagt ggcaggagg tcaagagaat gagtgaggac aggtgggtag    38640
gcagcccagg ccctagcaac agcaggagct caccccctcag tcactctagc caggactgaa    38700
atactttca ccctttcaag agagactagg aatctggatt tttatgtgaa atatcttgat    38760
tactaaatgt tgtcaacaga catgtcaaaa ggtaaaacta agtaagttca tgggagcagat    38820
tgactattca ggttatagaa ttaaggattc ttatccaaca cagataccaa ccaaaaagct    38880
gacgtataac atattaggag aaactatgtg cactgtcgaa acatcaacaa gggctaatg    38940
tctaaaaatag tctatattgg attccagttg aaacatgggg aaaggacatg aacaggcaac    39000
ttatgtcaat ggaaactcaa aaataaca agcatatata aaagcattct caaattcagt    39060
agtaaacaga cagatgcaaa taaaaagagg gaaactgctg ccgggcacag tggctcacac    39120
ctgtaatccc agcactttgg gaggccgagg cgggcggatc atgaagtcag gagatcgaga    39180
ccatcctggc taacatggtg aaaccccgtc tctactgaaa aaatacaaaaa ttagccggc    39240
gtagtggtgg gcaccagtag tcccagctac tcaggaggct gaggcaggag aatggcatga    39300
acccaggagg cggagattgc agtgagccga gaccatgcca ctgcactcca gcctgggcga    39360
ctgagtgaaa ctccatctca aaaatataa taataattat aattataata ataataaata    39420
tgcaaagatc agggccttgg gatgccgggg tgcagtggct cacgcctgta atcccaccac    39480
tttgggaggc caaggcgggc ggatcatgag gtcaagagat caagaccatc ctggccgaca    39540
cagtgaaacc cggtctctac taaagataca aaaattagcc gggcgtggtg gcgcatgcct    39600
gtaatcccag ctactcggag gctgaggcag gagaatcgct tgaacccagg aggtgggg    39660
cagtgaaacc cggtctctac taaagataca aaaaaatata tatatatata tatattatta    39720
tattatatat atatatcatca gagcctgggt aatccttgtg tgctgctggg gaaggtagtg    39780
gtgcagccaa ccttgacagt atcttcccagg tacttggtta tattaagtat aggcacacac    39840
aggagacatg cagtcctact cctgggtcta aatcccaaag aattctcaca caagtccata    39900
tctacaggag acgagtggga attcagcatt actgggagtg ggaatcaacc tgggtgtcca    39960
tctacaggag acagatgga caaaatgtgg tggatattaa gaccagaatc accaagtaac    40020
agagatgggt ggtgagtgac aatcctaaga tacagaataa aggctagaac atgatgccat    40080
tcatgtaaat taaaaataga tgcacacaaa gcagtatacg cgtgaccctt gaatagcaca    40140
ggtttgaact gcctgtgtcc acttacatgt ggattttctt ccacttctgc tacccccaag    40200
acagcaagac caaccctct tcttcctcct ccccctcagc ctactcaaca tgaagtgcac    40260
aaggatgaag acttttatga taatccaatt ccaaggaact aatgaaaagt atatttttctc    40320
ttccttatga tttttctttat ctctagctta cattattcta agaatatggt acataataca    40380
catcacacgc aaaataaatg ttaattgact gtttatatta tgggtaaggc ttccactcaa    40440
cagtaggctg tcagtagtta agttttgggga gtcaaaagtt atacacgatt ttctaactgt    40500
gcaggcaatc agttcccctg acccccctcat tgttcacggg tcaactgtat atacacaaaa    40560
gtattattta aacctcatta gaatagctgt ctatagggag aagagaatga gagtggggt    40620
aaacgga atg aacaaataaa ccaacaaatg cattaacaag caaaacaaca gaggggcttg    40680
catgggccag tgatgataaa gggctaagaa tgagaataata attaattcaa ttcctcacac    40740
ctgaggtcta aaaccaagga aaggggagggc caggcgtgga ggctcacgcc tgtaatccca    40800
```

                                    24

```
gcactttggg aggctgaggc gggcggatca caagattagg agtttgagat cagcctggcc   40860
aacacagtga aagcccatct ctacaaaaaa tacaagaatt acccaggtgt ggtggcacat   40920
gcctgtagtt agctactctg gaggctgagg caggagaatc acttgaaccc aggaggcgga   40980
ggttgcaggg agccgagatc acaccattgc actccagcct gggtgacaga gtaagactct   41040
gtctcaaaaa aataaaaaaa ataaaaaaac agagaaaggg aggaaactag atccaggctg   41100
actagataca gcctttagag ttagaaaaga tgatttgaca atctaagccc acactcagat   41160
tgaatgaaat tgaaaagcct ttcaaactaa aacatttaat tacaccatct gctgcagaca   41220
gaactcagac aactcaaaca ggtaatgtca gcgtggtgtt ttatatcacc accctcaaca   41280
cagaataaaa atcagctgca tgtgaagcag tgactagaat gaagaaaagg ctgcttctta   41340
cttccttcta gtggttcttt ccgaaaacat taataggcac cagccttatg catgtcaccc   41400
tgcagggaga catggggtat ataactatga cttactgttc attcctcaag gaattcccaa   41460
tcttgtggaa gattatacac aatgaggcaa caaaaactat ccaataaaac cacggaaaag   41520
aagccagtga caaagaagcc agtgatgaaa ggccctgtga gcagagctga tggccatttg   41580
gggaagaaag accaacatgg atggggtga tcaaggtggc tccgtgggaa agctggaaga   41640
gaagtggcag atctctgagc tggatgatgg gccactacca tctgtatatg gctaattaaa   41700
gaccatgtgt ggattttttt a ttcagctctt tcgtgtcatt cctgctatca gcacagaacc   41760
caatctcaac tttccagcta tattgagcta aacttctcac ctcatggaat ttgcagataa   41820
agttcaaaag gatccttgcc tttcaaaat aatttttgaat ggttgagtag tccctctgtg   41880
ctctctcact catcacatgc gaagcctgct gagggctgc catgctatgg ctttctccaa   41940
catcaggaaa gttctccac tcagtttcac cttaatacaa atgtgttctc tcttcagaga   42000
aggcaaaaaa atcatgacc atctgactgg gagaagtcat ttctaggtaa agttgatccatc   42060
tttttctgag gaacacagga ggaaaatctt acagaaaaga gttaacacag caggcctaag   42120
actgcttttt aaaataaata aataaataa taaataaata aataaataaa taaataaata   42180
aataaatgaa tgataggtc ttctgtattg gccaggctag tctcaaattc ctggcttcaa   42240
gagatcctcc cacctttgtc tcccacagtg ttgggattat agacatgagc catctggtct   42300
ggcccaagac tgttattcct aaaaagtctc ataaaaagca tggttaatcc ttggctggca   42360
cctgggaact tagatttcag aagggttccc accatccaac ctggaaagag ggactcactg   42420
tgcctaaatt attgtgtggt ttatgctgaa ctcctgcttt tcttcaggta gcgtggaatg   42480
tggtatgtgc tgggcaaagg gggcctgcat gaccagcccc caataaaaac cctgggtgtt   42540
gggtctctag tgagtttccc tggtagacag catttcacat gcgttgtcac agctccttcc   42600
tcggggagtt aagcacatac atcctgtgtg actgcactgg gagaggatgc tggtgaagctt   42660
gtgcctggct tccttggac ttggcccat gcacctttcc ctttgctgat tgtgctttgt   42720
atcctttcac tgtaataaat tacagccgtg agtacacac atgctgagtc ttccaagtga   42780
accaccagat ctgagcatgg tcctgggggc ccccaacaca gaaataaatt ataaaagacc   42840
aaggactggg catgtgggcc catgccggta atctcagcgc tttgggaggc cgaggcagga   42900
ggaccagtta agcccaaaag ttcaaagtta cagtgaccta tgactgcgcc aatgcactct   42960
aacctgggag acagagcaag accctgtccc caaaacaata aactaaacac atacttctgc   43020
cttccaagtg tcttaaaatt caatggaatg gtagaaacat ttttaaaaca ctaaatcaaa   43080
agaaacctga aaaacaagag tgccgatggc caactaaaag gtctaggaaa tttctgaaaa   43140
gtaaaaagta ctcagaacca gattacctga gcaaaccata gcccaatcaa agcttgggag   43200
gaggctgtta tgcagaagga aatggtaaca ggttcccagg aacagacttg taacagcaga   43260
tagaacagca gaggtagaac ctgacaaggt gattacctgg ggaactgcag tctgaatgac   43320
caggactgtt ggacccttcc cctcacatgg aatacacacg ccactcagca gcacaccaca   43380
gctcttcaac aatcacagga ggcacgctac gcctagtaag acaggaaaaa aggaattctc   43440
aaacttcgaa gatgaaacaa taaagaatca ccaagttttt attcagtatg atgaaacagg   43500
gacactgaat caacagaaca caaacccaag caaagataat tactagagca catagaagaa   43560
attattagat attcttggga agacctaagg ggacactata aagagcaagc agttggtatg   43620
tgcagatctt tgtgatatac caagaaataa aaacacagga tgaagaccag atagagaaata   43680
tgactactat ttgtgcaaaa aaggagaaat ggagaatctg attcatattt gcttgtatttt   43740
gcatgaagaa actttggaag gtacatagt aactaacaac aatggttacc tacttgtaag   43800
gcgagagaag taagaggaca ggaatgggtg gaacaccttt tgtgtcgga attgtgggt   43860
tcttggtctg tctggagtt gttccttctg atgtttggat gtgttcggag tttcttcctt   43920
ctggtggtgt cgtcagactc ctgactcagg agtgaagctg cagaccttcg cggcgagtgt   43980
tacagctctt aaggggggcg atctagagtt gttcgttcct cctggtgagt tcgtggtctc   44100
gctagcttca ggagtgaagc tgcagacctt cgagggtgt gttgcagctc atatgacagg   44160
tgcagaccca aagagtgagc agtaataaga acgcattcca aacatcaaaa ggacaaacct   44220
tcagcagcgc ggaatgcgac cgcagcacgt taccatctt ggctcgggca gcctgctttt   44280
attctcttat ctggccacac ccatatcctg ctgattggtc cattttacag agagccgact   44340
gctccatttt acagagaacc gattggtcca tttttcagag agctgattgg tccattttga   44400
cagagtgctg attggtgcgt ttacaatccc tgagctagac acagggtgct gactggtgta   44460
tttacaatcc cttagctaga cataaaggtt ctcaagtccc caccagactc aggagcccag   44520
ctggcttcac ccagtggatc cggcatcagt gccacaggtg gagctgcctg ccagtcccgc   44580
gccctgcgcc cgcactcctc agccctcttg tggtcgatgg gactgggcgc cgtggagcag   44640
ggggtggtgc tgtcagggag gctcgggccg cacaggagcc caggagtgg gggtggtca   44700
ggcatggcag gccgcaggtc atgagcgctg ccccgcaggg aggcagctaa ggcccagcga   44760
gaaatcgggc acagcagctg ctggcccagg tgctaagccc ctcactgcct ggggccggttg   44820
```

25

```
gggccggctg gccggccgct cccagtgcgg ggcccgccaa gcccacgccc accgggaact   44880
cacgctggcc cgcaagcacc gcgtacagcc ccggttcccg cccgcgcctc tccctccaca   44940
cctccctgca aagctgaggg agctggctcc agccttggcc agcccagaaa ggggctccca   45000
cagtgcagcg gtgggctgaa gggctcctca agcgcggcca gagtgggcac taaggctgag   45060
gaggcaccga gagcgagcga ggactgccag cacgctgtca cctctcactt tcatttatgc   45120
cttttaata cagtctggtt ttgaacactg attatcttac ctattttttt ttttttttt    45180
tgagatggag tcgctctctg tcgcccagac tggagtgcag tggtgccatc ctggctcact   45240
gcaagctccg cctcccgggt tcacaccatt ctcctgcctc aacctcctga gtagctgggga  45300
ctacaggcaa tcgccaccac gcccagctaa ttttttattt tattttttt ttagtagaag    45360
cggagtttca ccatgttagc cagatggtct caatctcctg acctcgtgat ccatccgcct   45420
cggcctccca aagtgctggg attacagacg tgagccactg cgccctgcct atcttaccta   45480
tttcaaaagt taaactttaa gaagtagaaa cccagcatag gccgggcgcg gtggctcacg   45540
taaccccagc actttgggag gccgaggcgg gcggatcacg aggtcaggag atcgagatca   45600
tcctggttaa cacagtgaaa ccccgtcgct actaaaaata caaaaaatta gccgggcgtg   45660
gtggtgggca ccggcagtcc tcgctactgg ggaggctgag gcaggagaat ggcgtgaacc   45720
tgggaggcag agcttgcagt gagccgagat agtgccattg ccttccagcc tgggcgacag   45780
agcgagactc cacctcaaaa aaaaaaaaaa aaaatagaga cccggaaagt taaaaatatg   45840
ataatcaata tttaaaaca ctcaagagat gggctaaaga gttgacggaa caaatctaaa    45900
tattagattg gtgacctgca aaaccagccc aaggaacatc ccagaatgca gcccataaag    45960
ataaagagag catttccgct gggcacagtg gtatggcagg ggaattgcct gagtccaaga   46020
gttgcaggtc acattgaacc acaccattgc actccaggcc tgggcaacac agccaggctgg   46080
tgtctcaaaa aaaaaaaaa ttaaattaaa aaagcacgaa tatttgagag aaaaaaaatgc  46140
ttatttcaag aaacatgaaa gataaataac gatattctaa ttcccaagta agaataattc   46200
cagaagcaga aaatagaata gaggcaagga aacactcaaa acttctccag tgccatagaa   46260
atgtgtatta atctttagaa tgaaacggac taccaaatgc tgagcaggaa gaacaaaaga   46320
gatccactct taagccagtg tggtgcccaa ggcgagtggc tcatgctgt aatcccagca    46380
ctttgggagg ccgaggcagg tggatcacct gaggtcagga gtttgagatc agtcaggcca   46440
acatggtgaa accctgtctg tactaaaaat acaaacatta gctgggtatg gtggtgcaca   46500
tctgtaatcc caactacttg ggaggctaag gcaggagaat cacttgaaac caggaggtgg   46560
aggttgtagt gagccgagat catgccacac tcccagcctg ggtgacagag caagattcca   46620
tctcaaaaaa aaaatccact cctagacaaa taatagttaa attttagaac accaaggaga   46680
aagaaaaaaa attgtaaagc ttcagagaaa ataaacatta actacaaaga aacgagagtc   46740
agacgcgtgc acttcttcct agataccagc agataaagca atatctccaa aattcagaag   46800
gtttttaacgt agaatcctat acccagtcaa gaatattcac atggaaaagt gaaataaaaa   46860
acattgttta aacatgcaag ggttcagaaa gtttaccatt cacagaatcc ctgaaaacaa   46920
aaccaaataa tcacttaagg actcattaag aaaacaaatg aaataaaaagc accaatgatg  46980
agtaaataat cagaaaaatt tacagtttac ctaaataact gtttatgcat aagtatgaa    47040
aacccaaaaa tttaatatgg gacagaatta aaatcatgat aagattcttt tttgctttac   47100
tcatggaggg ttcacataaa cagattatct tttaatgaca agagaaaaaa atgtttagat   47160
atgtgtgaaa aactaagggt accaaaacag tgcaaattca tttatcatca ggaaaatcca   47220
aattaaaacc acagtatcca ccagaataac taaaaggtaa aagacagaaa ttaccaagag   47280
ttggcaagaa tgtggagcaa ccacatatac ttctgtgggta aataagttgg tgcaaccggt   47340
actgaaaaact gtttgctagt atctactaaa accgagcaca tgcacagact acaaccaagc   47400
agttccactc ccagatacac actcaacaga aatgcacaca ctcactcaac aaaagacgtg   47460
tactagagtg ttcatgtact tactattcat aatagtccaa aaatgcaaac aaccaactgc   47520
caatcaaagt caaatgtata tctatattag ggatatatac aatggcatat acacagcaat   47580
gagaatgaaa tgaaccagcc cggcacagtg gttcatgcct gtaatctcag cactttgggc   47640
gggtaaggca ggcagatcac ttgaggtcag aaatttgaga ctagcctggc caacagggtt   47700
aaaacctgtc cccactaaaa acacaaaaat tagccgggca tagtgttggc cggaggctgta   47760
ttccagctac tcgggaggct gggttgggag aatcgtttga acccgaaagc cggaggtcgc   47820
agtgagccga gatcgtgcca ctgccactcca gcctgggcga tagagcaaga ctccgtctca   47880
aaaaaggaaa tcaaaaatat aaaataagat gacaggaata atccgcaaaa gatcagtaat   47940
caaaataaat ataaatgggc taaagctacc tattaaaaga caaagatttc acacccataa   48000
ggatagctat tatcaaaaaa agagagagaa taacagatgt tagcaaggat gtatggaaac   48060
tgaaattctc acgcattgct ggtgagaata taaaatggtt cagcctctgc ggaaaacact   48120
atgctggtc atcaaaaat taaaaataga agtactactt gatccaacaa ttctacttct     48180
gggtatatac ccaaataact gaaagcaggg tcttgaagag atatttgtac acccatgatc   48240
atggcagcat tattcataat agctatgatg tggaaccaac ataaatatcc tggaacaat    48300
atatggataa gcaaaatgtg gtgtatacat tcaatggaat attaattagc aataaaaatg   48360
aagaaaattc tgacacatgc tacaacatgg atgaaccttg agggcattac attaagacaa   48420
ataagccagt tataaaaaga caaatactat atgaggtact atattagata ctcatgcaag   48480
gtacctaaaa taggcaaatt catagagaca aaaagcagaa tggtggttgc caggggctgt    48540
ggtaatggat acagagcttc aattttgtaa gatgaaaaaa ttctggagat tggttgcata   48600
acaatgtgca cacacttaac actggggaac tgtaaactta aaagtagtaa atggtaaaaa   48660
taaaaataat aaataataaa ttttatgtta ttttaccaca atatttatta aaagacaaag   48720
attaactaat taaacaaaat ccagccataa gctaatggta agagtaacaa ttaaagaaga   48780
cacagaaaat tgaaaatcag tgactagaaa aagatattcc atataaatgc taacaaaaag   48840
```

26

```
caagtacagc aatataaaga gaatgaacaa aaaaaaaatt aaataagatg gctcgtttat   48900
tcccaaaagg tacaattcac caagaagata caagaattgt gaacctttaa gcacataaaa   48960
cagcttcaaa aatacaacat ttaaagaaaa atatatatta aacatagaaa tagtacaaaa   49020
acccctacaa gaatcataat gggagtcttc aatacaactc tccatatcaa aacttgaaac   49080
agagaaaaaa aataagttaa ggatgcagaa aacctgaatt accatcaata aacttgagat   49140
taatatagaa ctgtataccc aatatactaa ggttcaggg aacagtcgtg actgacagtg   49200
gactgcaaat taatctgttc ttaatctttg ttttttcttttc agcactgtgg cagaatagag   49260
atcctaaaaa ccttccagct acaaaacatc tttttaaaa tataaaaaa tacaaaaata   49320
actctgaaat caatagaaga cacatggtga aaccaaaatt ctagaataca gggagaataa   49380
aggcattttc agatattaca aaaacagaaa attgatcatt gctgaagtaa tttctaaaga   49440
atgtacttga gggagaagaa aaatgttcca aagaaaagta tctgtgatac agagaaggaat   49500
ggaaagtgaa gaaatggtaa acaggtagat aaagctaata aatgttgacc tagaaaataa   49560
caaaaacaat agcaataatg tctcgttgga agggttgaag taaaaataca attaaggcca   49620
aatgtgaggt aagtggaatg aaagaattag aagtccttgc cttgttcaca ggactgatta   49680
aataaatgag ccaggttttc cattcaaaca gttaaaactt gaacaaaata aactcaaatt   49740
aagtagaaag ataaaaaaca gaaattaatg tcatagaaaa ataaaaaatc aatagaatta   49800
atcaatataat cctggttaat aaaagctggt tctttgaaag gattaataaa ataatcatta   49860
agcaagtctg atcaaaaaaa aagagaaaag gtaccaaaaa aagtactgta tcagaaagag   49920
aacatacaga tacatacaga tatgtaagag tctgttttct tacaccagaa tactatatac   49980
aacattatgc tagcatatat taaatttcaa taatgttaat gattttctag gaaaacagaa   50040
aatattaaat ttactttgaa gaaacagaaa aactgagaaa aataaatgat catgaaaaaa   50100
atgaaaaggt aattaaatac tgatattaac tgcctaaaca acaccagcag cagcccaggc   50160
agtctgcagt caagttctgc caaacttgag ggaacagata attcttctat tccaaggcaat   50220
agaaaatgat ggaaagtttc ccaatttaat cagaagggac agcctgatcc ttgttatgaa   50280
cacagataaa aatggggtaa actatatgcc aaactcagat accaaaaccc taaataagat   50340
gctagcttat tgatgtgaac aatccaaaag tgcattttaa attagcccag ggttttagag   50400
aaagaaaatc tagcaatgtg accaccactt atgttaacaa ttttaagacg aaaatctaca   50460
tgatcatatc aatgcatgct acacaaaagc atttgggcaa aaaacccaac acccaccctt   50520
gactttttaa actcttagta attaggcata aacagaaatg tacttaatgt gatagaataac   50580
actcggtgaa gatacagagg gaatgctccc taaaaccaag cccaagacaa agattcctat   50640
ttaacctcaa tagtcaacac tgcagcgaga gtaatctatg gaagacaagg aaaaaagtaa   50700
aaacatgaga gacatctgtt gtttaacaga caataagatc aacctacttgg aagagcaaa   50760
cgaatcaagc gaaaaactat taaaactgag acaggcttta gtatggaggc tcagcttcag   50820
ctgtagtttg ggctaccaaa ttcaactcgc ttgcttggag agttaatcct gcaaagctaa   50880
tttctgttga ggtattagga ttgacaagcc tgtgctcctc cctcctcccc catcttcaac   50940
actgaaataa cacggtgttt ggaactggat aacagaatct tccaaaaaca aaaattgtcc   51000
tgaagggctg acttgtgccc ttactcaaaa aacactttat ctgctgcctg cagctcctac   51060
agttgctggt ggataagcct gccaaccagc tcggcgtaat tcttcctgca gagggcaagg   51120
aagagcactt tcacaggaaa atttttttcc gaactgtatg ccgcttatta cataaactta   51180
cgtgctggca aatggagctc cagcaaaata agatattcag agtcaaactt ccttaggaaa   51240
aaaaaaaaaa aaaagcaagc acataacact aatttccttg catgggcact ggggaaggag   51300
gtcgttactt ccgcacgccc gcaggtccgc accaccggga aacccacggg ggatcgcgcag   51360
tgcccccaggg ccttccaggt gcactgcgcc gcggcgcccc agctgacccg caccgcgcgc   51420
ccctagccct tccccctgtca ccccggccag gaagggcgg gagcgcggcg gacgccgagg   51480
gcgagggcct tctcggtcct ctgcaccacg cagcaccccc aaggcacaac agggaggggtg   51540
cggaggggctc ccgagaccca ggagccgggg cccggcctgc ccgcgcacct gtcccactgc   51600
cgggtccctg cgacaacctt cgggccccgga gggagggag cggccacctg ccgctgccac   51660
ctgcggcacc ggtcccaccg ctccgggccg gggccccgcc ggcgaagacg ccctcctgag   51720
ctgggggacag gacacgcgac gagggacccg gggccccccgc ggcgaagacg cagcacgcct   51780
tcccagaaag gcagtccgtc gccccccacga gccccggcag ggactgcgcg gaccccgcgc   51840
catccctttca gaccacgcgg ctgagggcca aagagccggc cggcggggcgg gctggcggcg   51900
cggctagtac tcaccggccc gctgtggtca gcgccgccgc aacccccagc ggccacggct   51960
ccgggccgctc actgatgctc aggagaggga cccgcgctcc gccggcgcct ccagccatcg   52020
ccgccagggg gcgagcgcga gccgcgcggg gctcgctggg agatgtagta cccggaccgc   52080
tctcatctcc ctatctgctc atctctctgt cgcacataat cgatgtttgg gcgtcccaag   52140
ccagatgtgg accccatttc cgcactctac actggaggtt ttctaagggt ggtgcccggga   52200
ccagcagctt cagcctcatc tgggaacttg agaaaatgca gattctccgt cgccaccagc   52260
ctattcggtt tttcctgcac taaaaccatg aagttgggac ccagcagtcc acattctcgc   52320
aagcccgtca agtgagttctg aggcggccctc cagtttgaga gctatgctca cggcctcacc   52380
tccgccccgc aaggagcccg gtcttgcctg tggcgctagc cgcacacgga cacctcatcc   52440
tgcggggccc gcccccccgc tgcacccctca ccgcccaacg cctcctccgg gatgcagcgg   52500
aggcgcctgg aagtcggcaa ggtcaacatc cccctctacc cctcacgcgt   52620
cctcctccag gggtgcctca tggccagggg ttagaaagag ccactgtgtt tcttgacatg   52740
gaagtggcct aagaccttaa tgaaaactgc aggagtggaa tgacagaacc tttggtcata   52800
cttgagggcg tgaagctcaa atgaggagga aggaaaggat gcaggagaa taaccaaccc   52860
```

27

```
tggcaagttg tggcgcccag gtagagggc gagcctaggc tagcggttct cgaccaggc     52920
cggtgttgcc cctcctcgcc gccccgcgta catttgggga ggtctggaga cattttggt     52980
tgtcatgatg cgggagttgc tactgttgcc taagtgggta gacacgaggg tgctcctcaa    53040
catcctacct gaaggacagg actgccccac aaggaagaat gatccggccc caaataagaa     53100
accctgggct ggtcagcaac aacccctttg ttctgagaag agaggaggaa agaataaaag     53160
aagtggggtg aagttttggt ttggtagagg aaacttgaag acattttcac tggaaagaaa     53220
gagaggaaga ggagggagat gtctgtaagg acgagcaaac cgggtgacag ctgattttcct    53280
catattgaag taatgagtcc tagttataat aaattcctaa taaaaaccca gtttatccct     53340
gcaataaact tgtctttttt ttttaaatat actgcttgat tctgtttgct aatattttat     53400
ttacaggctt tgcattgata tgcaaaaatg agatgggcaa taattttctt tttgaatgtc     53460
taatgttgtt tggtttcaga atcaatgtta tgctcacatc ataaaaaatt tggaaccgag     53520
gcaggaggag tgcttgaggc cagaagttcg agaccagtct aggaaacaca gtgagaccc     53580
cccatctcta caaaaaaaaa aaaagaaaaa aaaatgggca tgtttgcttt ttcctttac     53640
tctgaacaat ttaaggagca ttaaatatat ctattcttg aggtttgatc atttcccagt     53700
taaaaatgtt cctcccagcc tgatgctttc tttggggagg gtaaatcttt taaggctaga    53760
aaagtttctt ctgtgcaat tttattattt acattttaaa aattattcta gagttaattt      53820
tgataaagca tgtatttcatt aaaacaaatt atccttttt tccagatgct caagtgtatt     53880
tgcataaagt tgaggaaagt agtcttttgt gaatctttta acttctccca aatatcttat     53940
tttgtgtatt tttgcttctt tattttgta actttaaaa gtgtattttt ttttcaaaga      54000
atcagctctt aggtttatgt ttttggttat actggagctt tttctctctt cttttttaaa     54060
tatttttct cctttatttt ttagacgtat tttgctcaa cgtaatcgga agaaggtaaa       54120
ttagaatctt ttgttactat tgtgtttta ttctcctta tttctctgaa gtcctgcttt      54180
ataaatagta ccatgttatt tgtgcataaa tattcattg ttttatatc ttgggaattt      54240
tcccacttca tcatnaaatg accttcctg tctcatttaa tgtgttcaaa ctttgccctg     54300
aatttaactt tgtctgatat tttaccatcc tgctgaattt tgtttgttac ccaaacaac      54360
ctttgctgtt ttcgtctttt ctgaacctt tatttaggt aatcccttga attagagcac      54420
taagtttgc tttgtgatta aatctgaaaa tctttatctt gccatagatg agttgagccc      54480
tattcatgtg acagctatat tatgctgtt catagccctt ttggtcctt tttcactctt      54540
gcattgcata ttttgtgttt attgtgtttt gtgtttcctc tgataatttg gaaggtttgt     54600
atttttattc agggagttgc cttataatca tactccgcaa tacacatcgt cctcagtttc     54660
ttcagactgt ctgttaactc cctattctga ataaaaatga cattgtaatt tccctctttt     54720
ttctttaccc cttttctttct cctcacctaa tgtaaatgat tttatccttc tttagtattt    54780
gctttttaa ttaactacat ttataaatat ctttatcact tgattttaa atcagctttg      54840
aatgagatat ttggattcct agatataaaa gatgttaatt ataccatttc cacgttagta     54900
ggtttataaa atcatacatt ctgctgtgta accataatcc cacgtttgtt ttagttccac     54960
tcctacagtt aaaagattca gaagtattaat taacagttat tttgccatag ttttttcccc    55020
aacccatttt gtggtaagtt atgatcctgc tttagtttct taagaataat ttataacaaa     55080
gagtgtggtg gctcacgttt gtaatcccag cactttggga gacaagaggt agaaggatcg     55140
cttgaagcca gcagttcaag accacccctga gcaacatagt gagaccttgt ctctacaaaa     55200
aatttttaaaa tttagccaga cgtagtggcg tgtgcctata gtcccagcta ctcaggaggc    55260
tgaggcaaga ggattgctag agccccagaag tttgaggctg cagtgacctc tgattgtgcc     55320
actgcacccc agtctgggca agaaagtgag aacctatctc tttaaaataa tatttgttgc     55380
ttatgaaaat tatattcccct gagttttca tgtttaaaa tatttttatta ttgaagaatg     55440
taaaagtttg agtataaaatt cttgggttat aaccagaagt cagcctggtt tattttttcct    55500
cagaaatgag gtaattgccg gccggacacc gtggctcata cctgtaatcc caacactttg      55560
ggaggccgag acaggtggat cacgagtca ggagattgag accatcctgg ctaacatggt       55620
gaaacccccg ctctactaaa agtacaaaaa gttagctggg catggtggtg gacgccttga      55680
atcccagctc cccgggaggc tgaggcagga gaatcgcttg aacctgggag gaggagcttg      55740
cagagagctg agatcgcgcc actgcactcc agcctgggcg acagagtgag actccgtctc      55800
aaaaaaaaca aaaaaaaaca aagaagtgaa gtaattgcca tgatgctcca agaattatct       55860
cttttgtctat gaaatccaga aatctcactg ttatacattt tggaattatt tgaattttct      55920
aatatttcct gggacacaat agattgactc tatagatttta attttttttt ttttttttgag     55980
acagagtctc actgccaatct cagcttactg caacctctgc ctcacgggtt caagctaatcc      56040
tcctgcctca gcctcccaag tagctgggac tacaggcgcg tggcaccatg cctggctaat        56100
ttttgtcttt ttagtagaga caggtttca ccatgttggc caggctggtc ttgaacgcct         56160
aacctcaagt gatccacctg cctcagcctc ccaaagtgct gggattacag gcgtgagcca        56280
ccatgcccag cctcaattcc tctttctatc tggtaatttt tctgaagttg aaaacatttg         56340
ttctaatacg ttatttcagt gttcttctaa gatgtgtaaa gcaccctatt cccaggtcag        56400
ccccatcttt gctagtgagc tcggctggtt cttcacaaga gctctggttt tctcctgctt        56460
aatctcaagt acctctgtca gcctccacct ggtttatgat ttggagtttt ttggttttttg       56520
tttttttctt ttgacagatt cttactctgt cacccaggct ggagagcagt ggcataatct        56580
cagctcactg caactctgt ctcccaggtt tgagcgattc tcctgcctca ttagtagaga          56640
tagctgggat tacaggcgcg tgccaccaca cccggctaat ttttgtattt ttagtagaga         56700
tggggtttca ccatgttggc cagggtggtc ttgaactcct gacctcaggt aatccacctg         56760
cctcagcctc ccaaagtgct gagattacag gcgtgagcca ccgcgcctgg catggtttgg          56820
agttttaatc tgtagtttta ataaagatag tgcttatgtt tgtgtttctt atatttcttg        56880
```

28

```
gtactcttgg gtaatttgta agatccccat atctacacaa gaagtccatt ttcaattctt   56940
ttcttcagac tgtttatttt attttatttt attttatttt tatgtttgag atggagtctc   57000
gctgtgtcac ttctggaggc tggagtgcag tggcgcgatc tcaggtcact gcaacctccg   57060
tctcccgggt tcaagcaatt ctcctgcctc agcctcccga gtagctggga ttacaggcac   57120
ctgccacttt ttaatttttt tagagacaga gtctcgctgtt gttgaccagg ctggagtgca   57180
gtggtgcaat catggctgac tataacctcc aaatcctggg ctcaagtgat cctcctgcct   57240
cagcctcctg agtagctggg actacaggca catgccacca tgcccagtta atttcaattt   57300
ttttgtagag acagggtctc catatgttgc ccaggctggc ctcctactcc tggcctcaag   57360
taatcctcct acctcagcct cccaaattac taggattaca agcatgagcc accatgccca   57420
gccttgttct actacttttaa tttcatatgt taggtgacca tgtaattgat catccaaacc   57480
aggatactgt aagaatgaaa gaggctgaca gtagtatgat gctgggacta gcattgtgca   57540
ctgagattat ttctgggaaa gcaggagata cggtcaccct acttatagtg tgcttgtctt   57600
tggattgttg aatttggagt ttctatttgc aggcttattt caactgggca gccttgatcc   57660
gccctgccca gcaatgctac cgttctctcc accgggtctc tgggaccct tcagtcacta   57720
tacttagctc agttccccac cctcccactc cctaaaagcg taaccaggaa tcctgcctca   57780
ggtctactgc cgtcttccgt ggctgtttc agttcctatt acccagagtc aaactcccag   57840
cattccctac ctgattccag acttggagtc cagacgttta acctcttcag gccaactccc   57900
cactttgcat ttctgtccct atatcttagt ccatggagat acatttcatg tctttgagtc   57960
tacttacaaa gtaaattttg ctgtttttta attttttttt tgagatggag tcttgcccttg   58020
tcaaccaggc tgtggtgcaa tgacgccatc tcggctcact gcaacctccg cctcctgggt   58080
tcaagcgatt catctgcctc agcctcccaa gtagctgtga ttacagacag gcaccaccac   58140
gcccagctaa tttttttttat cttttagtag agacagggtt tcaccatgtt ggccaggctg   58200
gtcttgaatt cctgacctcg tgatctgccc atctcggcct cccaaagtgc tgagattaca   58260
ggcgtgagcc actgtgccca gccaattttg ctttttttat atttcattgc tatatgttta   58320
gaggataagt ttacagtgct atatgcattc ccaaatatta gaccaaaaaa atctccaaaa   58380
aattagaaag aaaatccaaa aaatctcaaa aaatacccaaa aagcaacaat ctcacagacc   58440
atactcactg acccccaata aaataaaatt agaaattaac cacaacttaa caaaataaag   58500
tactcaagtc agagaggaaa gaggaaataa acatcaaaat tacaaagtct aggcggtggc   58560
tcacgcctgt aatcccagca ctttgggagg ccaaggcggg cagatcacaa ggtcaggaat   58620
tcgagaccag cctggccaat atggtgaaac cccgttttcca ctaaaaatac aaaaattagc   58680
caggcatggt gatgtgtgcc tgtaatccag ccacttggga ggctgaggca ggagaatcac   58740
tgaacccagg gagacgaaga ttgcagtgag ccaaaatcgt gccactgcac ttcggcctgg   58800
gtgacaaagc gagactccat ctcaaaaaaa aaaaaattac aaactcttta gatagaaatt   58860
ttggtgtttt ttttttgagac ggagtctcac tctgtcgcag aggctggagt gcagtgggac   58920
tatgtcagct caccgcaacc tccatctcct ggattcaagc aattctcctg tctcagccttc   58980
ccaagtagct aggattacag gcgccaccca ccagacccag ctagttttta tattttttagt   59040
agagatggtg tttcaccatg ttggccaggc tggtctcaaa ctcctgacct caagtgatcc   59100
acctgcttca gcctcccaaa gtgctcagat tacaggcgtg agccaccgca ccccacctag   59160
atagaaattt caacatgagg ccgggcacaa tggctcaggc ctgtaatctc agcacttcag   59220
gaggctgagg cgtgggagga tcacttgggc ccaggagttc aggaccagca tgggtgacag   59280
agacagaccc tgtctctatt tatttgaaaa aaaaaaaaaa aaagagagag aagaaagaaat   59340
ttcaacatga aaagtatctc tcaaaccctt cgagatgttg gcaaaaagcg actcaaagga   59400
acacctgtaa tcccaacact ctgggagtcc gaatcaagtg gatcatgagg tcaggagatc   59460
gagaccatcc tggctaacat ggtgaaaccc tgtctctact aaaaatacaa aaaattagct   59520
aggcggcggtg gctcatgcct gtaatcccag cactttggga ggctgaggca ggtggatcac   59580
ctgaggtcag ggtttgaga ccagcctggc ctacatggtg aaacctcgtc tcttctacaa   59640
atacaaaaat tagctgggcg tggtggtgagg tgctgtaat cccagctact cagaggctga   59700
ggcaggagaa tcgcttgaac ccgggaggcg gaggttgcgg tgagccgaga tcgcaccact   59760
gcactccagc ctgggcaaca gcctgggtga cacagtgaga ctccatctca aaaaatacaa   59820
aaaattagga ggggtggtgtg gctcgcgcct gtagtcccag caggaggaggca ggctgaggca   59880
ggagaatgga tgaacctgg gaggaggagc ttgcagtgag ccagatccc accactgcac   59940
tccagcctgg gcgacagacc accatggcac   60000
gaaccctgat aataaagaaa ccaaatgttc aactccaaa gctcggacac tttaaagaaa   60060
taattaataa aggcagaagt taaaggggag atgataaagc aatttttttt gttggttttt   60120
ttgagatgga gtcttgctct gtcacccagg ctggatgca gtgatgcgat cttggctcac   60180
tgcaacctct gcctcccggg ttcaagcaat tctcctgcct cagcctcctg agtagctggt   60240
actacaggtg cgcgccacct ggcccacgca attttttgtat ttttattaga gacgggggttt   60300
caccatattt gttaggctgg tctcaaactc ctgatctcag gtaatctgcc cacctcggcc   60360
tctcaaagtg ctgggattac aggcaggcgc caccgccct ggcctaaagc aaaatattgg   60420
ttctgtgcaa aagtcaata aaaagagcaa agtttacaa actggagcca gcacccattc   60480
agctcagtgt gtctggagaa aaacaatct cgcttcagaa ttcatgatta cagaggttctt   60540
tttgcttcct aaaaatccta ctatgttgct gttgaccatt ctctctcttt ctctctctct   60600
tgctttctcct ccagaaaagc tattcagaca ttctcctctt tcctcaaacc tccaggagga   60660
cctcctccat ccttagcctc agctgctgac ctcacttcta atcattgaga aaccaggaga   60720
agcatttaag agtgaacctc cgcctccccg cacgggcaaa accacccacc cacagaattg   60780
tgccccaatt ctgcgtgctcc tcctctcacc atggatggac ggtccaggct ccgagccaaa   60840
                                                                     60900
```
                                      29

```
gccaggcctc ccctggagct ctggatccac cacctgcagc ttctcaggca gggccccagc   60960
agctcccctg ctcccttgta ccatcaatcc ctcccctcac tgggtcactc ccaacaatat   61020
atatatttag tgatgtttct cccatgtggt aaaatcactt agcctctctc ctcccccagc   61080
tactatccta tttgtttctt tccattctct gcaaaacttc tcaaagcatt gtgtctatgt   61140
gctgactcca tttatcttct ccgttctct gctgagtcct tcccacagac tctcaccccca  61200
gttactccat gaaatgacct ctgcactgcc acatccagtg gtgaatgttc agttcttaat   61260
tttattcagt ctttcagcag catttgacct ggccgatcac tccctcttct taaaaatact   61320
tttctcagcc aggcgtgatg gctcacacct gtaatcccaa cactttggga ggccaagcg    61380
ggaggatcat gagagcccag gagttcaaga tcagcctggg caacatggca agacctatc    61440
tctacaaaata ctaaaaagta gccagtgtga tggcatgcac ctgtagtccc actacttag   61500
gaggctgagg cagtaggatg acttgagcct gggaaatcaa ggctgcagtg agccatgatt   61560
gcaccactgc actccagcct gagtgacagc gagaccctgt ctcaaaaaga caaaatagga   61620
aactttttctc agcatattcc tctgattctc ctgctgcttc tgtctgcaca gattcagtct  61680
cctttgccgg ttcttcctca tcctcctgat ctcttgacct tgaagtgccc cagagtacag   61740
tcttttttttt ttttttttgag acgcagtctc gtctgtcacc caagctggag tgcaatggcg  61800
aggtctcagc tcatgcaacc tctgcctccc gggttcaagc gattctcctg cctcagcctc   61860
ccaagtagcc aggactacag gcacatgcca ccatgcccag caaattgttg tattttttagt  61920
agagacgggg ttttactata tcggcacgc tggtctcaaa ctcctgaact cgtgaaccac    61980
ccgcctcggc ctcccaaagt gctgagatta caggcatgag ccaccacacc cggcccagag   62040
tacagtcttt agacggcctc tctacctata cttgctcccc tcataaactc ctcctgcctc   62100
atggctttaa ataccatcgg tagactgatg actccataat ttctctttt tttttggaga   62160
cggagtctcg ctcagtcccc caggctggag tgcagtggcg cgatctcggc tcactgcaag   62220
ctcaacctgc caagttcaca ccattctcct acctcagcct ctccagtagc tgggactaca   62280
ggcacccggc accacgcctg gctaattttt ttgtattttt agtagagatg gggtttcacc   62340
atgttagcca ggatggtctc gatctcctga cctcgtgatc cgcccatctc ggcctcccaa   62400
agtgctggga ttataggtgt gagccaccgt gccagccga tgactcccat atttctatct   62460
cttgctgtgt gggagtttctc ctcagaactc catctcata aatccaactc tcataaatag   62520
tatctcaaat gggcaatatg ctcaaaagtc aattcctact tttctcccta aacttgcttt   62580
cctgcagtct ccaccatctt aatgtccaat ctaacattag gaggcaaaaa ctttgaagtc   62640
attcttgact cttctctatt acacaccta tccaatcttt ctgcagatcc agtcgacccc   62700
caaatccagt tagctctcat catctcccct gttacccect ggtccaggcc atcttcctct   62760
ctcacctgaa tcactgcagc attctcctca ctggtctctt tggttctgtt ttcactccac   62820
cttagcatag tctccacaga gacgcagag ggatcctttt aaagtgtaat tcccatcctg   62880
tccctgctct gctcaaaacc ctgtcgtgat tcccgtttta atctgtcaga ttaaaagcca   62940
gagtctttcc agtgacctac atgatctgcc tattatcacc tcccacttct ttccccttgc   63000
tcactccact ccagctctgc agctgtcctt tctgtttcct gaacagccca gattttgctt   63060
ctttagaacc tttgtatttg ctgtccctc tgtctggaat gtttttccag gaagtcacct   63120
ggctctctcc tgcacttcct tcctgaccac catgtttaaa aatcactcaa acacacttca   63180
ggccggacat ggtggctcac gcctgtaatc ccagcacttt gggaggccaa ggtgggtgga   63240
tcacctgagg tcaggagttc gagaccagcc tggccaacat ggtgaaactt cgtctctact   63300
acaaatacaa atagtagcca ggtgtagtgg cacacacctg taatctcagc tactcaggag   63360
gctgaggcag gagaatcgct tgaacccaga aggcagagga ggtgcagtga gccaagatca   63420
cgccacaaca ccccagcctg ggtgacagag caagacccca tctcaaaaa aaaaaaaagaa   63480
aaaaaaatca cacaaacaca cttctcttca tattcctttt ccaagttta ttttctccta   63540
gaatactttta cattgttttta atggaagttc tccgtttccc ccacactaga atggatactt   63600
cctgcaggta ggcactctag tcctcccatc caagtactaa ccaggctcaa ccctgcttag   63660
tgtttttattc attgctgttc tgttgattct cttttgttcc tcctcctagt gcttgaatga   63720
ctacttgtac ataataagca ttcaataaat attttgtgaa tgaatgactt gttgaatgaa   63780
ttaatctcag aaatgcagga ctggttctac attagaaaat tttcaaggt catctctcca   63840
tgtcgtaaca cattaagaga ggaaaattt gtactctaaa tcatttgata aaatacatac   63900
tgatttctgt tttcaaaaac tcttagtggc tgggcgaggt ggctcacatc tataatccca   63960
gcattttggg aggaccaggt gggcggatca cttgaggtca ggagtttgag accagcctgg   64020
ccatcatggt gaaaccctat ctctactgaa aatagaaaaa ttagccgggt gtggtggcgc   64080
atgcctgtag tcccagctac ctgggaggct gaggcaggag aatggcttga acccgggagg   64140
cggaggttgc agtgagccaa gatcgtgcca ttgcactcca gcctgtgtgt gtggtggcca   64200
ctccagctcc tctagtgaga gaggcaggaa aatgcgttga accccggagg   64260
ctccatctca aagaaaaact cttagtgagt ttaggaatcc aaggaagacc ctcaaactaa   64320
atagtaatc tagctaccag aagccttcag taaacctaa cactccatgg tgaaacatta   64380
gaaacattcc tactaaaaga caggctaaga atgcctgcaa tcttcacggc tagtccaaga   64440
agtcaaaaga aagaaatgag cgctgattta aaaaaataaa caaacaaaaa actaccgatg   64500
cagaggctgg cagcaaggac tgaaggactg tacagtactt gcctggagca ggcggatggc   64560
cacacccctg cgaagcctgc tcagctggct gggggacgct ccagtgtgtg aggtggcagga   64620
tgcagggtac ttcctctgcc agggagttgc actggggaa tcctcccccca ctcacacttt   64680
ggcagctgg gctttggaat gtgacttagc ttctgtcaaa gggtcaatcc acccctttgat   64740
atatgatgca aaggcgaaca tatgatgcaa aggtgagaga acagcccaaa ttaggacttt   64800
taccacagct gtggaggtgg acacgcacag tgtgtgggcc tggccagact tttcatgctc   64860
aaaggtggtg gttgttcttc ctacttcttg tccctccagg gcttccttg cctgtgtgct   64920
```

30

```
gaacctgctt cttttaattt tttttaactt ttttaaattt ttaattgttt taattaaaac   64980
aaattttgaa aactgtctga acctgctttt gaaccctgct atgatttgaa tgtttgtccc   65040
ctgccaaact gattttgaaa cttaatctcc aaagtggcaa tattgagatg gggctttaag   65100
cagtgactgg atcatgagag ctctgacctc atgagtggat taatggatta atgagttgtc   65160
atgggagtgg catcagtggc tttataagag gaagaattaa gacctgagct agcatgtcyg   65220
cccctteace atttgatatc ttacactgcc taggggctct gcagagagtc cccaccaaca   65280
agaaggctct caccagatac agctcctcaa ccttgtactt ctcagcctct gtaactgtaa   65340
gaaataaatg cctttttctt atgaattacc cagtttcaga tattctgtta taaacaatag   65400
aaaacgaact aaggcaaact ctcatgattc tactgccatg ccattccaat aaactccctt   65460
tatgcttaag agagccagag ttggccaggc gagaccatcc tggctaacac   65520
ttgggaggcc gaggcaggtg gatcacaagg tcaggagatc gagaccatcc tggctaacac   65580
ggtgaaaccc cgtctctact aaaaatacaa aaaaattagc tgggcgtggt ggtg   65640
tgtagtccca gctactcggg aggctgaggc aggagagaa tggcgtggac ccaggaggcg   65700
gagcttgcag tgagtcgaga tcgtgccact gcactccagc ctgggtgaca gaatgagact   65760
ccgtctcaaa aaaaaagaga gccagagttt atttctgttg cttgcaacca agaaatctgg   65820
ctggtgacct gaagtttcca taaataatag caatttaaag actcatttcca agccaggcaa   65880
tgcctagcct tgtgtagtct ttgtggtaat acattcattc attcatttgt tcaaccaact   65940
gtgctccgaa gactaagaat acaaaaatgg gggccgggtg tggtggctca cacctataat   66000
cctagcactt tgggaggccg aggcaggtag atcacctgag gtcaggagtt cgagaccaac   66060
ctggccaaaa tggtgaaacc cctactctac taaaaataca aaaaattagc tgggggtggt   66120
ggcggaaccc tgtaatccca gctactcgtg agactgaggc aggagaatca cttgaacccg   66180
ggaggcagag gttgcagtga gccgagatcg caccactgca ctccagcctg ggcaacaaga   66240
gcgaaactcc acctcgaaaa aaaaaaaaaa aaaaaaaagg ggccgggtgt ggtggcgcgtg   66300
gctcacgcct gtaatcccag cactctggga ggccaaggcg ggagaattac gaggtcagca   66360
gatcgagacc agcctgacca acatggtgaa accccatctc tactaaaaat acaaaaatta   66420
tccgggcgtg gtggcgcaca cctctagtcc cagctacttg ggaggctgag gcaggagaat   66480
cgcttgaacc cgggaggcag aggttgcagt gagccgagat catgccactg cactccagcc   66540
tgggtgacag agtgagactc cgtctcaaaa aaaaataaa aaaaaaaaaa gaattcaaaa   66600
attgtagagt tatagtgtgc ttctagttta gttgagagga catctgtcct tcaaggaagg   66660
ctagaatcta taccctgagt ccttactgaa atcaatccag cagtcaaaac atgggaccaa   66720
cgatcacagc agtaagatag gaagagcacc tttgtacatt tagctcatgt tgagataagc   66780
cactgacaga gctgaaggaa gctcacagtt ctgggttcca tcctttggca tttaaaaaga   66840
aaagtgctaa gaaaattcgg ttggtcacgg tggctcacgc ctgtaatccc acactttggca   66900
gaggccaagg caggcagatc acgaggtcag gagttcgaaa ccagcctggc caacatggtg   66960
aaacccccgtc tctactaaaa acagaaaaat tagccgggcg tggtggcgca tgcctataat   67020
cccagctact caggaggctg aggcaggaga attgcttgaa cccgggaggc ggaggttgca   67080
gcgagtgaga gcaggccact gcactccagc ctgggagaca gagcaagact ctgtctcaaa   67140
aaaaaaaaag aaaaaaagaa agaaaggaaa aaaagaaaga aaaaaaaga aaaagaaaa   67200
ttcaggccag gccaggcctg gtggctcaca cctgtaatcc caacactttg gagggctacga   67260
gcgagacggt gccttagccc aggagtttga gaccagcctg agcaacatag cgagaccctg   67320
tctctataaa aaaaaatttt ttttggcca gacgcagtgg ctcacgcctg taatcccagc   67380
actttgggag gccgaggcag gtggatcacg agtcaggag atggagacca tcctggctaa   67440
cacggtgaaa ccccatctct actaaaaaat acaaaaaatt aaccgggcgt ggtggcgggc   67500
gcctgtagtc ccagctactc gggaggctga ggcaggagaa tggcgtgaac ccgggaggcg   67560
gagcttgcag tgagccgaga ttgcgccact gcactccagc ctgggcgaca gagcgagact   67620
gtctcaaaaa aaaaaaaaaa aaaaaaaaat taattgtcag gtgtgctggc atgcagctgt   67680
agtcctaact actcggggag ctgaggcagg aagatcgctt gagcccagga gttcaaggct   67740
gcagtaatag tgcctctcac tctacctgtg gtgacaactg gacctctct caaaaagaaa   67800
gaaaaaggag aaagaagaaa agaaagaaag aaagagaaga aaggaaggaa gaaagaaga   67860
aaaagaaaag gaaggaagga agaagaaaa aaaagaaaga aagaaaagag agagaagttc   67920
aaagaccaaa gggtcaggat cccaaaatag ttttttatgt ttatttattt atttacttat   67980
ttattttttga gacagtatgg ctctgtcgcc caggctggag tgcagtgatg cgattgcggc   68040
tcactgcagc ctccaaactg ggctcaggtg gcctcccac ctcagcctcc cgagtagctg   68100
ggaccacagg cgcgtgccac catgcccagc taattttta attttttgta gagatgaggt   68160
ctctatatgc tgcccaggct ggtctcgagc tcctgggctt aagccatcca ccgcctcggc   68220
cctcccaaag tgctgggatt acagaatgca gccaccgcgc ctaatcggtt ggtttgtttg   68280
tttattgacg gggtctcgct gctgcccagg ctggagtgac agtggctgtt cacaggtgca   68340
gtcctggacg attgcatcag ctcttgggct ctagcgatcc tccagagtag ctgcaggctg   68400
gattccagcc gcgccaccgc gcggggctca gaatgggttt ttatattggg ggttatgctg   68460
ccacctagag gatatatgta gtaccgaact gtgtcgccag ggaggctgag gtctgcagtga   68520
gccaagatga tgccagggcga ctccagcctg ggtgacagag caagatttca tctcaaaaa   68580
aaaaaaaaaa aaaaaaaaaa aagaattgaa agtaaggtct tgaaagata tttgtgcctg   68640
tatgtgtcata gcagtattaa ctttgaccca ctagctaaaa cacaaaagca acatgtgtct   68700
gtcagcaggt gaacggataa acaaaatgtg gtatatatgt acaattgaat attattcaguc   68760
ctttaaaaag gaataaaagg ctgatgcggg gggctcacgc ctgtaatcct aacactttgg   68820
gagactgagg tgggtggatc acccgagggtt aggagtttga gaacagcctg gccaacatgg   68880
tgaaacttca tctctactaa aaatactaaa attagccgag catggtggca cttgtctgta   68940
```
                                      31

```
atccaagcta ctgggaggc taaggcagga gaattgcttg aactcaggag ccgaggttg    69000
cagtgagcta agatggcacc actgcactcc agcctgggca acagagtgag actccatctc    69060
aaaacaaaca aacaaaaaat tattatttcc aaagaaacaa gaccctgggt ccatttccca    69120
gcccacacct gatgttgact cacaacacac agcctggttt gctatgagcc tgcttcattt    69180
aattgtcacc ttaacttcac atcaccctca agtcctggaa taactcttg ctgaccttg    69240
tgtgctgagc catctccatg tcgctcaacg tgcagtccct ctcactcgac tgagtcaata    69300
gccagacgtg gtctgactgc agggtcatcc ttggtggctt aggctgactc gggcatgaca    69360
gggtgctctg agacctcacc gcataaggc ttggctcccca ataaactcta tataatattc    69420
ctctttccca ggatttcctc ctctacctcc tcaagtccca ctgctctgca aagaccaaaa    69480
atattatgtg gtctgggtgt gtgtagcttc gcactgtctt ctcgtgacag tgccctcaac    69540
tttaaacaga gtcttttact gcagatccca agaacagcca caccctctc tcccacccac    69600
gctgcagagt cccagctccc tccttacac cccacgacgc agcctcctct ctcagaaccc    69660
tccagcacaca cccaggtaat tatagcaccc aggtaacta tgtagatgga gtccctggaa    69720
catgtggata gtgcccctg ggagtatgca aaagcaacat tgctggcacc tgcagagaac    69780
agggtgacat ccaggaatca ggacatgggc ctctgggagg tagggatgtg gcaggcagg    69840
ctgccaaaca ttggtagagc aaggccacag gatctttctg accttccttc caaacagagg    69900
ctcctgtact ggtgatccct gtgttgattg accatccct tcctgggggt cgtggtctct    69960
gtccccagttg cccggacttc tgtgagtgtc ctactgaggt cctttcatg agaagcatgc    70020
tgtccttcca cctgctggga gcaagagtga caacttcaat actataatag cagtggcacc    70080
cagagaagaa gaaagatgaa gtggcaagaa aaacagcgtt ccaagcagga gtttttctat    70140
aaaaacaaaa acgtttacaa gcaaactttt tataaaggc tagatagtaa atattttagg    70200
ctttgagagc cacatagact tgtttgcagg gactcaatgt cgctattgta gtttgaaagc    70260
agccatcagg gttatgtaaa tgagtgagtc tgattttgtt tcagcaaaat tttatttacc    70320
aaaacagaca atgagtgggc tggatttggc ccatgatcct tagtttgcca actcctgctt    70380
tgggctcacc cagatctgat tttgaattct ggctctgcta ctggttagct gcaggagcct    70440
ggaaggctct ctgagcctgt ttcctcatct gtaaaattaa agcaataatt tctaacactc    70500
aagagtgtta cctcacgcct gtaatcccag cactttgggag gctgaggcag gcggatcacc    70560
tgaggtcaga agttcaagac cagcgtggcc aacgtggcaa aaccctgtct ctactaaaaa    70620
atacaaaaag tagccgggca tgtggcgcg catctgtaat cccagctact tgggaggctg    70680
aggcagggat actgctagaa cctgggaggt ggagcgtgca gtgagtggag atcacacctc    70740
cacactccag cctggccgac agagcgagac tccatctcaa aaaaaaaaaa aaaaagagtg    70800
ttagaaggtt ttgagataat gaataaaaga tgccttgtgt atactaagta ttcaacaact    70860
gatagctgca ttggtcttaat tataacagtt tagaagcgat tgagtcaaca aatgctggat    70920
ttgtcaggga ggacttccta tcaggaggta gatcttgggc tgagtcctga agcaaagata    70980
ggcattggat agaggagttg agagaacacc ctaggactgt tattattatt attcgacacg    71040
gagtctcttg ctctgtcacc caggctggag tgcagtggcg cgatctcggc tcactgcaac    71100
ctctgcctcc caggttcaag cgattctcct gcctcctaag tagctgagac tacaggtgtg    71160
tgccaccaca cccggctaat ttttatattt ttagtagaga cagagtttca ccatgttggc    71220
catgctggtc tcgaactcct gacttcaggt gatccacccg cctcagcctc ccaaagtgct    71280
ggaataacag atgtgagcca ccgcacccag cccagaacca tttttcaatc cttggctctg    71340
cctttatta gctgcaagat ctcaggcaat ttatttaacc tctccaaaga ctcattttct    71400
cattcacaaa atgaggcaaa taataatatc tactatccca ggttgtcatg agaattaaat    71460
gcaacatgac atttaatgaa atagaaagtc ccttggacat taactggcta aagtatgtac    71520
tcgacaagga tatcattta ggtggatact cagcatcca gaactgatgc tcacaatgga    71580
atatcattga aacgcattaa aattcatttt aaatgattgt aggtagtgag gcaattgaaa    71640
gaagaagaca agaggactga ttataatgct tcaggctcac tagtctcctt ttaggagagga    71700
aaaacaattt caagttaaat tttaggctct agattttac ccctgctgct cattagaatc    71760
acccagattg atgaaatcag aagccaatcg aggctgtgtt tttcatctcc ttggatctcg    71820
ctgttgtggg gattaagttt ttgaaaaagt acatctaaca ggtgatcgaa aatgatagtg    71880
atattattgc agtgatgtgtc aattattgtt ttattattat actgaaagag gcttcagttt    71940
tctgatccat aaagtgaggg aattgcatga gaccattgct aagattcctt ctagctctgt    72000
tttttgttt ttgtttttta gacagagtct ctgtcgcccca ggctggagtg caatggcatg    72060
gaagtagctg ggactacagg cacacaccac catgcccagc taacttttat attttttaata    72120
atcttggtc actgcaacct ccgctcccg ggttcaaatg atcctcctgt ctcagcctcc    72180
gaagtagctg ggactacagg cacacaccac catgcccagc taacttttat attttaata    72240
gaggtggggt ttcaccatat tggtcaggct ggtctcaaac tcctgacctc aggtgatcca    72300
cccgcctcgg cctcccaaca tgctgggatt acaggcatga gccactgtgc ccaacccctt    72360
ctagctttct tgatcactga ttctcagggtt ctctgctgaa atatatttga gacatcctgg    72420
ataaagatc atgcaaagac tcccaatatg gtattaataa ttgattctgg aggcttagct    72480
actcctgatg gattagacat gactcaactg cctctcttat gtgtacaaca caacacaca    72540
accaagaaag gttattctgg cattccattt attcagttta tttacagccc ttacttccag    72600
cagcacgtta aagatatggc caggggccgg tgcagtggct caagtctgta atcccaggac    72660
tttgggaggc caagtgggcg ggatcacaag gtcaggagtt tgagaatctg gcaattcttc    72720
agacttagaa gcaaccagct cgataacaca gtcttgtgtg ggctctccct ctgtccctcc    72780
ctcgcttccc tcatttctca tccctgcccc tgagactgtc caccttcaca tagccctgcc    72840
atgagacctt catctcaggc tttgctttct gggtaactg aggctaaaca ctgagtggcc    72900
ctaaaagagg attgggattt ggaagttaga ttattcacca gagaacagac tttgctgatg    72960
atcaggccca ggttgtaatt gttgaaaaaa agagaggatg catagtctta tctcatctcc    72960
```

                        32

```
tagtcaaagt caacaccatg ataaataaga gtcaaatcct gagatgtgaa ttgggggacat    73020
ttgagtggtt aaccctgaga agcttgcacc ttcagacccc tcaataaccc tgctccccag    73080
agaaggctgg acattgacct cagcacaggc aggagccctg caagatgcca tttgtcctac    73140
taaagatgga cccctccact ctgtttctag gtaaataacc aaagtcaagt ctccacacag    73200
cctgagcaag aaagtcagag cctgctacag gagaaaatac cacactggcc aaaggattca    73260
ctagccctgg ccactgtgtg tgggaggaac cagggaatca tgtgtgggag tcaatgttga    73320
agctgtttga ctgggggtgg ggtggaatat aagcctggcc ctggggagtt tttcccgttt    73380
gagggccttt acccacaact caagatccag tgctatagca ggagatccca gagctagtcc    73440
taacagatgg tcaggattga acttggccta gagtaaaatg aggaggatag tgccagaact    73500
ttctcaacat actattgagg aagaggtcag aaggcttaag gaggtagtgt aactgggaaag    73560
gggtcctgat ccagacccca ggagaggggt cttggacctt gcataagaaa gagttcgaga    73620
cgagtccacc cagtaaagtg aaagcaattt tattaaagaa gaaacagaaa gagttcccaga    73680
ccatagagca gcgacatggg ctgcttaact gagtgttctt atgattattt cttgattcta    73740
tgctaaacaa agggtggatt atttgtgagg tttccaggaa aggggcaggg atttcccaga    73800
actgatggat ccccccactt ttagaccata tagagtaact tcctgacgtt gccatggcgt    73860
ttgtaaactg tcatggccct ggagggaatg tcttttagca tgttaatgta ttataatgtg    73920
tataatgagc agtgaggacg gccagaggtc gctttcatca ccatcttggt tttggtgggt    73980
tttggcccgg ttcttttatca catcctgttt tatgagcagg gtctttatga cctataactt    74040
ctcctgccga cctcctatct cctcctgtga ctaagaatgc agcctagcag gtctcagcct    74100
cattttacca tggagtcgct ctgattccaa tgcctctgac agcaggaatg ttggaattga    74160
attactatgc aagacctgag aagccattgg aggacacagc cttcattagg acactggcat    74220
ctgtgacagg ctgggtggtg gtaattgtct gttggccagt gtggactgtg ggagatgcta    74280
ctactgtaag atatgacaag gtttctcttc aaacaggctg atccgcttct tattctctaa    74340
ttccaagtac cacccccgc ctttcttctc cttttccttc tttctgattt tactacatgc    74400
ccaggcatgc tacggcccca gctcacattc ctttccttat ttaaaaatgg actgggctgt    74460
ggcgcggtgg ctcatgcctg taatcccagc actttgggag gccgaggcgg gcggatcatg    74520
aggtcaggag atcgagacca tcctggctaa cacggtgaaa ccccgtctct actaaaaatg    74580
caaaaacatt agccaggcgt ggttgcaggt gcctgcagtc ccagcggctc aggaggctga    74640
ggcaggagaa tggcgtgaac ctgggaggtg gaggttgcaa tgagccgaga ttgtgccact    74700
gcactccagc ctgggtgaca gagcgagact ccgtctcaaa aaaaaaaaaa aaaaaaaaaa    74760
tagctgggca tgttggcgca tgcctgtaat accagctact ctgagggctg aggcaagaga    74820
atcgcttgaa cccagtaggc ggaagttgca gtgagccgag atcttgcac tgcactccag    74880
cctggtgaca gagtgagact ctgtctcaaa aaaaaaaaaa agaaaaaaaa agacagaaag    74940
aaagagcaca gacagagtca caggtatttg cagtaggaag ctgtcaggtt agagtgcacg    75000
gaaatagaaa gtatattttta cacttacagc acatcttcgt ttgattagcc acatttaaaa    75060
tactgaatag caacgtgtgg ctatttagta ttcactaaaa tcttggacag tgcaagtcta    75120
aagaatcctt gatccgtccg gcatggtggc tcacgcctt aatcccagca ctttgggagg    75180
ccaagtgga aggatcactt aaggtcagga gttcgagacc agcctgccca acatggtgaa    75240
acctcgtctc tactaataat acaaaaatta ttagccgggc atggtggtgc atgcctgtaa    75300
tcccaggtac ttgggaggct gaggcaggag aatagcttga atccaggagg cgctgcagtg    75360
agccgagatc atgccatgcc actactgcac tccagcctgg gcaacaagat gccaacatgg    75420
caaaaaaaaa aaaaaaattg ttgggcgtgg tggctcacgc ctgtaatccc agcactttgg    75480
gaggctgagg ggggtggatc acctgggttc tggagttcga gaccagcctg gccaacatgg    75540
tgaaaccca tctctactaa aaatacaaaa attagctggg cgtggtggtg ggcacctgaa    75600
atctcagcta ctcaggaggc tgaggcagga gaattctttg aacccaggag gcagaggttg    75660
cagtgagcca agatcgcgcc tctgcactcc atctgggtg gcaagcaag actatgctc    75720
aaaaaaaaa aaaaaaatac ttgattgtct ggacattctg cagaacatca tatggagaca    75780
ctatgttgac gacatcatgc tgattgtaag caagaaatgg caagtgttcc agaaacacag    75840
tcaagacaca tacatgccag aaggtgagat ataaatctca ctaagattca gtggcctgcc    75900
acactgggtga catttttaaa cctgctagat gtttgtgtag aaaaggattt aacctttgccc    75960
aaagaggggt ctggcctttg tccccagcta ctggacataa tctcttttaaa ctcttgaaat    76020
atcattcctg atagaagtat tttttgtttg actaggggcc ttgggccagc cagatagcaa    76080
caatgtgatc tgggttgggg gctttggatc aggtggcatc agtgtgacct cctgagtggc    76140
tagagactag aatcaaccac atgggcagac aacccagctt acatgatgga attccaataa    76200
agacttttga cacaagggct tgggtaagct ttcctggttg gcaatgctct atactgggaa    76260
acccattctg actccatagg gagaggacaa ctggatattc tcatttggta cctccctaga    76320
ctttgcccta tgcattttttc ccttgtctga ttattattat tattatgaga tggaatctcg    76380
ctctgtcacc caggctggag tgcagtggaa tgatctcaac ctcactgcaac ctctgcctcc    76440
ccggttcaag cgattttcct gtctcggcct cccgagtagc tgggactaca gatgcatacc    76500
accacaccg gctaatttttt ttgtatttttt agtagagacg gggtttcacg ttagccagga    76560
tggtctcgat ctcctgacct catgttccgc ctgcctcggc ctctcaaagt gctaggaata    76620
catgtgtgag ccaccgcgcc cagccccctt ggctgattat taaagtgtat ccttgagctg    76680
tagtaaatta taaccgtgaa tataacagct tttagtgagt tttgtgacaa cttctagcaa    76740
attatcaaac ctaaggatac ccttggggac ccctgaactt gcagttggtg tcagaaacaaa    76800
gggtgctcat gtgtgtacca tgccctctaa ttttgtagtt aattaacttt cacaactttta    76860
ttattaccgc ttacactcaa tgtttattca catttatcca cataccactt attctagtgc    76920
cttgcatcaa agactttcta tctcatgtac tttattctgc ttgaagtaaa tcctttagga    76980
```

33

```
tattcttttt ttttttttaaa ctttgcacat acatactttt attttttatt tattttttaat   77040
tttgttattt ttgtgggtac gtagtagata tatgtatta tggagtacat gagatgtttt       77100
gatacaggca tgcaatgtga aataagcaca tcatggagaa tggggtatcc atcctctcaa      77160
gcaatttatc cttcaagtta caaacaatcc aattacactc tttaagttat tttaaaatgt      77220
acatttaatt ttgtattgac tagagtcact ctgttgtgct atcaaatata attttttttt      77280
tttttgaagc agagtctcac tcagtggccc agactgaaag tgcagtggca caagctcggc     77340
tcacttcaat ctctgcctcc ctggttcaag cgaatctcct gcctcagcct cccacatagc     77400
tgggattaca ggcacacacc accatgccca gctaattttt atattttttt agtagagacg     77460
ggttttcgcc atgttggcca ggctggtctt gaactcctgg cctcaaatga tctgaccacc     77520
tcagcctccc aaagtgctag gattacaggc atgagccacc acacctggcc aaaatagaat     77580
attcttagt gaggtctgct ggtgacaatt tttttctttt ttttgagact gagtctcgct      77640
gttgtcagct tgggctggag tgcaatagca cgatctcagc tcactgcaac ctccacctcc     77700
cggattccag caattctcct gcctcagcct cccaagtagc tgagagatta caggcaccca     77760
ccaccacacg cggctaattt ttgtattttt agtagaaatg ggggtttcacc gtgttggcca    77820
ggctggtctc gaactcctga cctcaggtga tccacccacc ttggcctccc aaagtgctgg     77880
gattacaagc atgagccacc acgcacagcc aatttttttcc gttttttgtct gaaatcttat   77940
tttgtgtcat ctttgaaata tatttttgat ggatataaaa ttgttggttg atagttatta     78000
tcattattat tattattttg agacagggtc tcactctgtt gcctatgctg gggtgtagta     78060
atgtgatctc ggttcactgc agactctgacc tcctagggct caggtgatct tcccacctca    78120
gcctccctag tagctgggac tacagatgca tgccaccata ccaactaat tttttcattt      78180
tttgtagaga tgaggctttg ccacatttcc caggctggtc tctaactcct gagctctagc     78240
aatccaccca ccttggcctt acaaagtgct gggccatgac tagccagcag tttacttttta    78300
tagcatattg aatatttaat atgaatcttc tggcatccac tgtaactgtt taaaaaatca     78360
gctgtttact tggcactctt ttttttttttt tttttttga gacagagtct tgccctgtcg    78420
cccaggctgg agtgcagtgg cgtgatcttg gctcactgca agctctgcct cccgggttca     78480
cgccattctc ctgcctcagc ctccggagta gctgggacta aaggcgcccg ccaccacgcc     78540
cggctgattt ttttgtattt ttcgtagagt tggggtttca ccgtgttagc caggatggtc     78600
tcgatctcct gacctcgtga tctgtccgcc tcggcctccc aaagtgctgg gattataggc     78660
gtgagccacc gcgcccagcc tctttttttt ttttttttag acggagtctt actctgtcat     78720
ctaggctggt gtacagtggc gtgatctcag ctcagtgcaa cctccacctc ctgcctcagc     78780
ctgccaaata gctgggatta caggtgcgta ccatcacgcc cggctaattt ttgtattttc     78840
agtagagatg gggtttcacc atgttagaca ggctggtctc gaactcctgg cctcaagtga     78900
tctgcctgcc ccagcctccc aaagattaca ggcatgagcc accgcacccg gccaagtagc     78960
actcctttga aggtaatctg cttccctac ccctagcaat ttttaacaat ttttcttcat     79020
ttttatttcc tgaagttttg ttattaataa tctgtgtgca gatttctttg tatttcttttt   79080
gtttgcagtt catagtgatt cttgaattag tgtgttggtt tctgttatca ccacaggaaa    79140
attgtcagcc gttagttttt caaatatttc cttgctaaat tctctctttct ccccttctgg  79200
tacaattgat ttgattaaaa ctaaaaccag ggccgggtgc agtgactcat gcctgtaatc     79260
ccaacactt gagagggctga ggcaggtgga tcacgaggtc tcaggagttc aagaccagcc     79320
tggccaatat ggtgaaaccc cgtctctact aaaaatacaa aaattaccag gcatggtggc     79380
acacatttgt agtcaggagg ctgaggcagg agaattgctt gaatccagga ggtggaggtt     79440
gcagtgagct gagatcccac cactgcagtc tggcctgggc gacagagtga gatgagaatc     79500
tgtctcgaaa aaaaaagtta tgaatgtttg ataaactata tttgttagaa tgtttgttgt     79560
agaatacatat tcattgattt ttaaacaatg ttagataaaa ccattcactg gatttgtgat    79620
aattaactta ctgattttac ctcactgatt tgttgtaatt aatacaactg gtataaaaag     79680
actgtgacga ggccgggcat ggtggctccc gcctataatc ccagcactt cggaggctga     79740
ggcaggcgga tcacctgagg tcaggagttc aagaccagcc tgaccaacat ggtgaaaccc    79800
catctttact aaaaatacaa aattagccgg tcgtggtggt gcatgcctgt aatcccagct    79860
cttcgggagg ctgtggcagg agaatcactt gaacccggga ggtggaggtt gcagtgagcc     79920
gatatcgcgc cattgcactc cagcctgggc aacaagagcg aaactccgtc taaaaaaaaa    79980
aaagaaaaaa aacacatgaa acaaaacaac actgtgacgg ttcccaaaaa ttaggagcat     80040
aattaaagga actcctgata aaaattaatt ttatcttaca tgtaaactaa aatgactttta   80100
tgaagttaat tcagaaatac aatgcagggt attagtttgc cacacgtgcg tattcagcct    80160
aatgtaatat tcttgttatt tttaaattct tctttaactt ttactcatat gtggatcatc   80220
aaatttcaaa agattaaatg acaatactct tagcagcaag cttccctaag catataaaca     80280
ttttaatggg tgatgattca gaaggtaccc gaagaatatg tactgccaga tactcattcac    80340
ccccatatac ctgcccgaca gacatcccat tttgggaccc tggatataatg tgtgggtgga    80400
gagaaagata ggagaaagtg gtataagcaa atggctttg agtctgattg acacggcgattg    80460
aaatcctgtc tctacctctt aacagcctca tgatcctaca taagttaccc cgatcctcag     80520
ggccacatct gtaaattggg ggttgcgatg gcagccatct cacagggtct cttttcgggg     80580
aagggcagga attatggatt aagtgagcta gtaattgtaa agcacttaat acaaggaggg     80640
cgcataataa gtacttcata aataatgacg gccattatca tgactgaggt gtatgtcagct    80700
gtcggggatt acggcgactt cagaatttct gtggggtcagg gctcaaaggc agcaaatcac     80760
actggaaggc gaggtgaggc actgcttctg cacagactgc ttagctggag acaggatgga    80820
aggcttagag gagatttaga ggaacttaga gtcctccgcc tccaactctg tgggatctgg    80880
tcccgtgcca gagacattca gggggatttct cgcactcctcc cctccccctac gtccctccg   80940
cccatccaa ctaaccacac aacacataca aacagataca cc tgcgaggttc tgcacgctgg  81000
```

                                    34

```
aagggaacag gagaagggcg ctgcgctttc ttgctgatgc cctgtacttg ggcccctggt    81060
agacacagcc acttgtcccc tcagcctgca gagaaatccc acgtagaccg cgcccgggtc    81120
cttggcttca gccaatctcc ctttggtggg ggtgggatgc acgatccaag gtttttattg    81180
ctacagacag cggggtgtgg tccgccaaga acacagattg gctcccgagg gcatctcgga    81240
tccctggtgg ggcgccgctc agcctccgg tgcaggcccg gccgaggcca ggaggaagcg    81300
gccagaccgc gtccattcgg cgccagctca cttccggacgt ccggagcctc tgccagcgct    81360
gcttccgtcc agtgcggcctg gacgcgctgt ccttaactgg agaaaggctt caccttgaaa    81420
tccaggcttc atccctagtt agcgtgtgac cttgagcagt tgactttatt tttcagtgcc    81480
tagttttcca gataccagga ctgactccaa ggactattac tcatctggag ggtttagcac    81540
agtaccgtcg catagtaaat ttccatgtca gttttggtta cctttcatgc acttgacaaa    81600
atgccatgct ctgaaacgaa ataggcacat ctttttttt ttttttttta aggagtcttc    81660
ctctcgccca ggctggagtg cagtggcgcg atcttggctc actgcaacct ccacctcccg    81720
tgttcgagat tctcctgcct cagcctcctg attactgagg actacaggca tgccacgacg    81780
cccagttaat ttttgtattt ttagtagaga cgggggttcg ccatcttggc caggctggtc    81840
taactcctga cctcaggtga tctgactgcc tcagcctctc aaagtgttgg gattacaggc    81900
ataagccact gcatctggcc agaaatgaaa taagtaaact ttttaacctg ctctaacaat    81960
atagtgaaaa gaccatatta ttattagacg aggttaaggg atttgcctat ttggatccct    82020
agttatagtc ttaaacttgg acattcttgt agaaagtaaa aagtttcctc ttcaaagttc    82080
ccttcttgt taaagaatac atcataagtg ttagaagtaa tagtttattt taaagactaa    82140
ctttcttcaa gcctccttgc tttgtgctaa taactctttg ttaagcccta tcctatgtaa    82200
ctgttggaca tgctcacagg cacgttccag ttcacagcct atgccccttc atctttcct     82260
aatgttattg cttccttaaa gttttaggct attagcaaat ctctcttctt cgttcttcct    82320
tgaggtcccg ctccagccaa tggatgcagg acatagcagt gaggacgacc caaatgcgta    82380
agggataaat atgtttgctt ttccttttgtt caggtgtgct ctcgacatcg ttccatctgc    82440
gattgagcac ccttctgca gaaagtaaag attgccttgc tggagatctt ttgtctccgt    82500
gctgactttt cttcgtggca ccgattatct attctaaca attttggtat ttctaacatt    82560
ctgaacaatc ttgggctagt tgtctcttct gggcctgttt cccatccgt cacatgataa    82620
acttcattgg tttaaaaacc ccagcgaaca tttattgagt tactattacc ttcctgccct    82680
ccccaacccc aaccccaggg agcagttaca acctcagccg ctgagcgcac tcgccgggtg    82740
ttaagaagca ccaaagacag ggaggcttga ttgattttgc tttgggagta gagggtcaga    82800
agattcacag gaaaatggca tttgagcaag gatgattcac tggagctagc ttttaaaatac    82860
tggcgaggct tttatgttgc agtcccttac aaagttgagc attcgcaggg actgcaactcc    82920
gaaataagcc cgcttcccct tttcattcgc taatgatcca gggagctgct ggttccggcat    82980
gcggcaggtt gtgcttttc ctaatcaggg ttctgcatcg cctcgaaccc gcaggccgtg    83040
gcgggttctc ctgaggaagc agggactggg gtgcagggtg aagctgctcg tgccggccag    83100
cgcctgtgag caaaactcaa acggaggagc aggaggggtc gagctggagc gtggcagggt    83160
tgaccctgcc ttttagaagg gcacaatttg aagggtaccc aggggccgga agccggggac    83220
ctaaggcccg cccgttcca gctgctggga gggctcccgc cccagggagt tagttttgca    83280
gagactgggt ctgcagcgct ccaccggggg ccggcgacag acgccacaaa acagctgcag    83340
gaacggtggc tcgctccagg cacccagggc ccgggaaaga ggcgcgggta gacgcgggga    83400
gtcacgtggg cgatgcgggc gtgcgcccct gcacccgcgg gaggggatg ggggaaaaggg    83460
gcggggccgg cgcttgacct cccgtgaagc ctagcgcggg gaaggaccgg aactccgggc    83520
gggcggcttg ttgataatat ggcggctgga gctgcctggg catcccgagg aggcggtggg    83580
gcccactccc ggaagaaggg tcccttttcg cgctagtgca gggcgccctc cgcgccagt    83640
agtccgggcc ggttgctgaa tgagggggagc cgggccctcc ccgcgccagt ccccccgcac    83700
cctccgtccc gacccgggcc ccgccatgtc cttcttccgg cggaaagta gctgaggagg    83760
cgccgccggg gagtcaggcc gggcctcagg ggcggcggtg gggcaggtgg gcctgcgagg    83820
gctttcccca aggcggcagc aaggccttca gcgagcctcg acctcggcgc agatgccccc    83880
tgagtgcctt gctctgctcc gggactcttc tgggagggag aagtggcct tcttgcgcga    83940
ggtcagagga gtattgtcgc gctggttcag aagcgattgc taaagcccat agaagttcct    84000
gcctgtttg ttaagaacag ttcttaggtg ggggttagtt tttttgtgtt tctttgagga    84060
ccgtgatca agatcaagga aatctcttta gaacctatt atggaagtct gaagtttcca    84120
aatgttgagg gttttatgtc taaaagcaac acgtgaaaaa attgttttct tcacccagtg    84180
ctgtcttcca atttcctctt tggggggagg ggtagttact gctgttacta aaataaaatt    84240
acttattgct aaagttcccc aacagtaaga ccactactt tgatgacttt ggcaagtttg    84300
ctaactactg gaaccctaac ttacaaacga actacttaca tttttgattt ccagttgtat    84360
ttcattaaaa atacatatcc gaagtgagca agtatgggtc tgtggacagc agtgattttt    84420
cctgtcaatt cctgttgctt cagataaaat gtaccagaca gaggccgggc gcggtggctc    84480
acgcctgtaa tcccagcact tggggaggct tggcgggtgg atcacctgag acgggagagt    84540
caagaccage ctgaccaaca tggagaaacc ccgtgtctac taaaaataca aaattagcca    84600
gggtggtggc gcatgcctgt aatgccaact acttgggagg ctgaggcagg agaatcgctt    84660
gaacctggga ggcggaggtt gcagtgagcc gagatcacac cattgcactc cagcctgggc    84720
aaaaagaggc aaactccgtc tcaaaaaaaa agtaccagac agaaatggt tttgttttct    84780
ttttttgttt tgagacggag tttcgctctt gttgcccagg ctgagtgca atggcgcgat    84900
ctcagtctcg gctcactgca acctctgtct cccaggttta atcgattctc ctgcctcagc    85020
```

                                    . 35

```
ctcccaagta gctgggatta cccatgcccc accatgcccg gctaattttt gtatttttag   85080
tagaaacggg gcttcaccat gttaggctgg tcttgaaccc ctgacctcaa gtgggcctcc   85140
cacctcggcc tcccaaagtg ccaggattac aggcatgagc caccgcggcc agccagaaat   85200
gggttttgga aaaagcacta aacaaaatcg aacttggttt catatgacag ctctgctgct   85260
aactgtaaca ggggcagaca agttaaccta cttttctgtc ttctgtcagc tgagaattag   85320
atgattccca aaggcccatt gaactctgaa tgactttaaa tacttcttct taagtgggta   85380
cacggttttg gtaactgatg ccaggtgatg aatgcatgaa agtgcttaat gaatgaaacc   85440
ggtaaaatag taggaggaag ctttattggt aaggcaggg tataactaat agctctctaa   85500
tttattggta ttgaagtggt taacttttgt tttttaagg ggggaaaaca ttctaagaat   85560
aatgaggcaa actgcatatt gcacaagaga ctgtcgtctc tattcaacaa ataccttttg   85620
agtgtccaga gtctgccagg tgctgtgcta actggggtaa acagtaaata aataagcaaa   85680
atgtccctgc acccatggag cttattgtct actgggtaag acagataata aataagcaaa   85740
caagttgagt aaagtgttat taaagattaa gaggagaaat gcattttggt tagatgttag   85800
aggactcagc aggtgacctc gaaacttaga gggacttaac cagtaggagg taactagaga   85860
ggccagggaa tcgcatgttc aaaggccagg aggcaagaaa gagcatggtg ccctccaaga   85920
gaggaaagaa gaggttggtg actggagcat agatgtaggc aagtgttggg tgattgagag   85980
ctctacgggc catggttagg ttttattcct aatgccggaa tgccaaacat ggtggttcat   86040
atctgtaatc ccagtatttt aggaggccga ggcaggaata tagcttgaac ccaggagttc   86100
aagaccagcc tgacgaacat gagacctgta caaaacattt aaaaaattgc tgggtatgat   86160
ggtgcacacc tgtggtccca gctactcagg aggctgaggc agaaggatca cttgagccta   86220
ggaggtggag gctacaatga gccatatttg agtcactaca ctccagcctg gatgacaaag   86280
tgagaccatg tgtcaaacaa aatacagaaa gaattataat ttaaaattt gaaagaggag   86340
tgatctgaac ttatatctta aaaagatcat tctagggcat ggtggctcat gcctgtaatc   86400
aagggctttg ggaggctgag acaggaggat caactgggag cagttcgaga ttcaactgta   86460
cagcatagag agactccatc tctacaaaaa gaaaaaataa atagctgggt gttgtggagtt   86520
attcaggagg ctgaagcaga aagtcactt gagccacgga gtttgaggct gcagtaagct   86580
atgatcccac cactgcaaca cagtgagatc ttgtctcaaa aaaaaaaaa aatcattcta   86640
ggtgcttttt ggaggctgga tgtggtaaga gtagaagctg gagatggtcc tgttagggat   86700
tcgattcaga cttaaaatac catcaatgca ttgagtccca aatttacatc actacgttgg   86760
atccttgccc ctgaatccag actggtaatt ccaactttag gttcagtttg tatctctacc   86820
tgaccaatat agaggtgtcc agtcttttgg cttcctcagg ccacattgga agaagaattg   86940
tcttgagcca cacatgagt acactaacgc taacaatgac agatgagcta aaaaaaaatc   87000
gcaaaactta taatgtttta agaaagtta cgaatttgtg ttgggcacat tcagagccat   87060
cctgggccgc gggatggaca agcttaatcc agtagataccc ttcaacttac aatatctaaa   87120
attttatgcc agatttagtc atttaaacc tgctcatcag ttttttctcaa gaagtagtat   87180
tttggctttt tttcttttct ttttttttgag atggagtttc gctcttatcg ttcaagctgg   87240
agtgcagtgg cggatcttgg ctcactgcaa cctccgcctc ctgggttcaa gtgattctcc   87300
tgcctcagcc tcgcaagtag ctggaattac aggcatgcgc caccatgacc agctaatttt   87360
tggagacagg gtttcaccat gttggtcagg ctggttttgt actcctgacc tcaggtgatc   87420
tgcctgcctc ggcctcccaa aggctgggat tacaggcgtg agccacccgc ccggcctgca   87480
tttttggatt tttagttgct cagcccaaaa ctttagtaca tcttgaacc tcttctttcc   87540
tcctactcta tatctgatcc atcagcaaat acaccctggt tacctcacac atatcgaaat   87600
cctaccacgt ctcaccatct gtgacaatta acaccctgtt ctaggcagtc atctggcatt   87660
agattgagtg gttaaggatg tcctctaagg agatgacatt caaatcttag cttaaatgtc   87720
aagagggagc tggttttata aagattgagg agcaggatt attttgccat aggctttccat   87780
ttggtttcca ttccattctt gatacttatg gtatatattc aaaacaaatg cacagaaaca   87840
gacccaggta tattgggaat ttcggatata gagttcctag ttgggaaaag atagactgat   87900
ctgtaaatga tgctagttat ccatcatctg gcaaaaaata atttcctgcc tcctctcata   87960
tatctcagat caacagactt ttttcttgtt taa gggccaaatc ataaatatt tagggcttccc   88020
agaccatatg gttctctgta cactctcctt tatccttgaa gccatagaca atatgtaaac   88080
aaatgggcat ggctgtgtca cgataaaact ttacttacaa aaactggtag tgggccaggtt   88140
taggcatgcc cagcactttg ggaggctaag gcagatggat cacttgggggt caggagttttg   88200
agaccagcct ggccaacatg gtgaaaccct gtctctacta aaaatacaaa aaatagctgg   88260
gcatggtggt gggtgtctat aattccagct actctggagg ctaagacaca agaatcactt   88320
gaacccagga ggcagaggtt gcagtgagct gagatagcac cactgcactc cagccagggt   88380
gacggagtct aaagcaaaa caaaacaaaa ggtagtgggt tgtatttggc ccatgggctg   88440
tagtttgcca atccctgatg cagaaacaaa ttccaggtaa ataagagcct ggaatgttaa   88500
aaaaacaaaa cttgaagtca tgtagaagaa caggtaggg gaacaatcct gatctcagga   88560
taggaaggga tattgcttaa aataagacac aggaaaatat aatccatgtt gtgtaaattt   88620
gactacgtta aaacttaaaa ctttcgccaa gcgcggtggc tcacgcctgt aataccagta   88680
ctttgggagg ccgaggtgag cagatcacca ggtcaggaga ttgagaccat cctggctaac   88740
acggtgaaac cccgtctcta ctaaaaatac aaaacattag ccgggcgtgg tggcgggcgc   88800
ctgtagtccc agctacttgg gaggctgagg caggagaatg gcctgaaccc gggaggcgga   88860
gcttgcagtg agtcgagatc gcgccactgc actccagcct gggcgacaga gtgagattcc   88920
gtctcaaaaa aacaaaacaa aacaaagcaa aaaactaaa actttcatac aataaagtat   88980
acctaagata cttctagaag agaagtttta catccaggac gtgtatggaa tttctgcaag   89040
```

36

```
taataagtaa aagacaaggg acatgaagag gcagttcaca aaagaggaag ccaaaatgac   89100
caataaacat gaaaggatgt ttaacctcaa aggaaacaag gaaatgaatt aaaaacatca   89160
aatgccattt caaaactagt aagttggcaa aattaaaaat accaaggatg agaatatgaa   89220
gcatggcat atgagtgcat ggaatggtac agtcactttc attaaaaatg cacataattt   89280
gtttttattt tattttttg agacagtcta tgtcgcccag gctagaatgc agtggcatga   89340
tctcggctca ccacaatctc tgcctcctgg gttcaagcaa ttctcctgcc tcagcctcct   89400
gagtagctgg gattacaggc acatgccaca acgcccggtt aagtttgta ttttagtag   89460
agacagggtt ttgccatgtt ggccaggctg gtctcgaact cctgacctca ggtgagctgc   89520
ttcccaaagt gctgggatta gaggcgtgag ccaatgctcc tggctgaaaa aaatgcacat   89580
aatttgttac ctagcaattc catgtctaga ggcttatcct agagaaatct tgcttatat   89640
gcataggaag acgtgtacta gaatgttcac tagttgaatg tttaagtgaa aattaggaaa   89700
taaagtaaat gttcattaac aggaaaatga attataaaa caattaagta   89760
gctaaaatga ataaactaga gctgcgtgaa tgaactagaa ctggttcaat agtcatgtca   89820
gattattgaa tgaatacagg tcagatatgt atagagtgtc atttgtgtaa ttaatttttt   89880
ttttttttt gagatggagt ctcactctgt tgcccaggct ggagtgcagt ggcgtgatct   89940
cagctcactg caacctccac ctcctgggtt aaagtgattc tcctgcctca gcctcccgag   90000
tagttgggat tacaggcatg caccacgctg ttgaactcct gacctcaagt gttccaccca   90060
acttggcctc ccaaagtgct aggattacag gcgtgagcca ccgtgctcag ccatttgcgt   90120
gattttaaa gatgtgcaga ataatgccat taaaaaaat acacatacat gtatatatat   90180
acacgttttg ctgggtgtgt tggctcacac ctgtaatccc agcactttgg gaggccgagg   90240
caggaggatc acttgagccc aggtagtacaa gactagcctg ggcagagatag caagacccca   90300
tctcaacaac agaaaggata attaggtatg gtggcatgag aggatcactt gagcccaggga   90360
gttcgagtgt tatcaggcca ctgcactcta gcctggacaa caaagcaaga ccgtgtctca   90420
aaaaaataaa aataaaaagt atttgtatgt ggtcatggct aaaaaacgta catggaagga   90480
aaatgtcttt atttatttat ttattttttt ttttttaaga cagagtcttg ctctgtcacc   90540
caggctgggg tacagtggtg taatctcagc tcaccgcaat ctcggcctcc cgggttcaag   90600
cgattccttct gcctcagcct tctaagtagc tgggactaca ggtacccgcc accacccct   90660
gctaattctt gtgtttcag tagagacagg gtttcaccat gttggccaag ctgtctctcga   90720
actcctgacc ttaagtgagc caccgccctt ggcctcccaa agtcctggga ttacaggtgt   90780
gagccacctgc gcttggccag gaaatatcta atttagtaag tatttatatc tgggaaaagga   90840
agggtcaggt ggtgattcat aggaactcta aagtctatgt ataatactta gggggacaga   90900
aggaaaataaa gcaaaatgct gatatttgat tgttgagttg tgtatatgtt agaagtataa   90960
cataggagat ctgattgata gtaggagaat gtttttaggt ggtaaaagtg gaaccgtggt   91020
ggtttgtttt ggcagtagaa tcagttggtc atagtttgta tgtggaaggt aataaacaga   91080
ccatgttaag gatgacttcc ggaattttgg tctgagtagt gggtggatga cagtgtcatt   91140
catgagggaa gatgaagact gaggtaggaa caggtttggg agaagatgac atgttccctt   91200
ttagacaagt ggaattatgg aagatggcag gtaggtggtt agctaatga atttgagata   91260
aaagatttag gatggagata taaatttagg agtaacagcg tatctatggt attgtaagcc   91320
ttaagaatgg gtaggatcag ccaggaaata cagatgtata tgcagaagag aggagtcaag   91380
gaagccagaa caagttaatg tttaaagtga gtgatgtagt ccatgggcag atgctgctga   91440
gagggctgca aacaccagtg accctacaac attttaaat gtcgtcttcc tgacagcagt   91500
gatcagtaccc tgcaacgatc ttatttattt ttttcatgtt agtctccaca cacttgaatg   91560
tagacttttt gaaggcaaaa tcattgcctt ttctgagctg ggagcatgtc tggcacatac   91620
caagcactca acagttgatg tattgacttc atccagatac tctgaggggcg agttattcc   91680
tgctactagc ctttcacctt tcaatgttta agagcacaaa tacagagatg ggcacgtttt   91740
ggcatttctt atttttgataa cctttttccb gtcagatttt tcaatgttga aaaaaaaaaa   91800
caagaaaaga gggttaaaaa tagtcttatg tcagatcctg tgatggaatt cacacttggc   91860
ttaagctgct gggcacccttc ctatcttgga tgtcatattg tgtcatatta gcttatctac   91920
atctttttatt ttttttagacg gagtctctctt ttgtctccca ggctggagta cagtgtcgcg   91980
actctcggctc actgcaacct ccgctcctcg ggttcaagca attctctgcc tcagcctcccg   92040
aagtagctgg gcttacccag tccgcagccc ggttgggagc agttattttg aaaaaaaaaa   92100
caagaaaaga gggttaaaaa tagtcttatg tcagatcctg tgatggaatt cacacttggc   92160
atctttttatt ttttttagacg gagtctcgct ttgtctccca ggctggagta cagtgtcgcg   92220
agatgtgggg tttcaccatg tcggccacc tggtctgga ctactgacct caggtgatcc   92280
accggccttg gcatccaaa gtgctgggat tacaggcgtg agccaccgtg cctggcccag   92340
acaaattatt atactctgag tgttagaggc ttaggatgtt ttccttgat gctatggag   92400
gaataagtaa taagatatga tacacaacca aagacctttc ttcactatgc ttctagtagc   92460
tagtactatg gatgacacat ggtaataata ttggttagca tttgtcctca atttactgtg   92520
ctagttactc ttctaagccc cttacaggta tatattttt ttcatcaata atcctctaag   92580
gtagtttttta ttattgctt aatttataa atcaagaaaa ttaagaccca gagaagtaag   92640
taacttgtcc aagatcacat ggcttataag tggtagagcc agaatttgac cccagatgtt   92700
gtgactacat tgtctctcca taagcaggtt caactctttt gactggatgc tgttccaagg   92760
tcacttcctt agagaagtct ttgctgacaa ctaccctcct gtgccctcct ccaaggctgt   92820
ccattgttct agaactttga atactcatct tagaataaag ctggtctaat atctttttat   92880
ttataggaatg gatctctgac tgcaaaagtt ggtcataatt atcttttat gttctagtga   92940
aaggcaaaga acaagagaag acctcagatg tgaagtccat taaaggtaag ttctgcccttt   93000
ggcagtccac tgcattaaaa agtgatgtgc tttgcattttg tgagttcttt aatcctgtta   93060
```

                                    37

```
tactctctct tttggcatta atcatttctg ccttatttta taattactta tgattttgat    93120
ttatttccct ctttaacctg tataatgctt taacatctag catataataa gtaggctttt    93180
tttttttttt tttttttgga gacggagtct tgctctgtta cccaggctgg agtgcagtgg    93240
cgcgatcttg gctcactgca agctctgtct cccgggttca caccattctc ctggcctcagc   93300
ctccccagca gctgggacta caggtgcacg gcgccacgcc tggctaattt tttgtatttt    93360
ttagtagaga cagagtttca ccatgttagc cagtatggtc tcgatctcct gacctttgtga   93420
tccgcccgcc tcggcctccc aaagtgctgg gattacaagc gtgagccacc gcacccggcc    93480
gtaagtaggc tttttttacc ttaattttat ttttttgaga tggagtcttg ctcttatccc    93540
caggctggag tgcagtggtg ccatctcggc tcactgcagc atccacctcc cgggttcaag    93600
cgattctcct gcctcagcct cccgagtagc tgggattaca ggtggccgcc accatgccca    93660
gctaattttt gtatttttag tagagacagg gtttcaccgt gttggccagg ccagtctcaa    93720
actcctgacc tcaagtgatc cactcgcctt ggcctcccaa agtcctgaga ttacaggcgt    93780
gagccaccat gcctggccat aagtaggctt ttactgagcc ttgtgtgtat tggctatcct    93840
agtgattaca gtgaaccagc gccctttctta ttaatcacac atttaatTgt tccctaaaag   93900
tgattagttc actttattta tttagtaaga caaaaaatga agaatactct taactgagca    93960
gtctgttaac tgtaggaaag cactgacact tataaggctt agttttctgt catttatcca    94020
gaagtatggt tgattacagt ttttactttt ttattgaat gaacaacctt aatttaaaat    94080
atattttgtt tattttttgt tgggatcgat acattgtcct tgtttataga tctagagcatg    94140
ctttttaaag atgctgtatt actcactgat tttatttgtc cagtgtacag agatttaacg    94200
gggaaaatta taatggaaat tgtttccata gtcattacat attaatttca tcaatttatt    94260
tccataaaat ctgtagattg ctacttattt agatttttcc ttcaaatgtt tttatgttgt    94320
attgcttgca ctgatatttt attctatatg ctcaatttgc tggagaagaa gactaattat    94380
aacttaggca agttgtaaaa ttagggaaaa aagtaaggta ccttacagcc tagttcaatc    94440
atttcttatg taaagccagt tagattccac attagttcaa actgccttct ttgagcaaaa    94500
cttgattggc agtgataaag gcttaaagcc cttctcaagc agagacctgt aaagactaga    94560
tctgactgta gtagaaggaa ggaacttaga tgtttcaggc agtgagaaca ccagtcttcc    94620
actctaaact ttgccactaa cagtatgacc ttgggaagtt gtaactttct tcagattctt    94680
catttgttga atgggggat tggcctagct aatttctaaa tctctactgg gctaaaaaat    94740
tctgtgctta tactctgatt atgaagtaca taatctgtgc ttaacattca ctgacttatc    94800
cttaggataa tacagaagca gtacaagaaa cagcccctca agatgtttgc agtctggtta   94860
gaaagacaaa cttatacaca gaacagtagc aaatagacca aaataataat agctggccatt   94920
tatagaacac ttcttctgtt ctgggcatta gacaaaaact gactataacg gtgaacaaaa    94980
aagacttagg tcctgccctc attgaactta cagattagta ggggagagga acattaatca    95040
agtaattcca cagatggctt agcctagatt ggtgagtgatg gaagtaaaga gatgtgaacg    95100
gacttgaaaa aaaattcgga ggcaaaatgg atagaagttt attattgatt aaatatgagg    95160
tgtgagagag agggatattt aagattgata cctaccttct ggcttgccta acagaaccaa    95220
aacagaaaat tatatgttca gttttgttat gttgggtggg aggtgctttt ggtgtcattca    95280
tttatatatg ttatatatgt tattttatat gcatagtaat tttaagtgct gagttttaaa    95340
ccaaaggtta gagagtgatt ttttagagtc tagcaaaacct aagttgaaat cctgcctgtt    95400
gaaatggctg tttactagct cattaaccta gggacaaagta ttcaacttgt tttcattttt    95460
gtcttcatct ctaaaatgag gaaaatatgg tcttacaaga ttgtcctgag agatagatga    95520
aataatatcc aaaaaaaaaa aaggtacata gagaaactcg tatagtgcct ggtatatagt    95580
aggtcctcca ttggtagcta tcattatcta gttttaacat agcctcagt ttgttgaatt    95640
agtcaaactg agtgaagcac tgcaaggaat tcagaggaat ttgagatcaa caatgggaat    95700
ctgaagttta gggaagactt catggcaatg acacttacct tgtataaaag ttgaagaata    95760
agaaagattt gaatgagaga ttctttctct tctccctacc agcccagctt attttatgag    95820
gatatattgg gcaaaggggc cttcagacaa gtagagggag attttttacag aaagatttgag  95880
atgaaggtat agaaggctgt aaagaccaga aaagagaagt gagacagagg aagcaggaag    95940
ccactgtagg tttttgagca agatattgat gctgtaagta tggtgtttat gaaaggttag    96000
tctggaagag atttgcagga tggagacccc gggagtttt ttgttataat acagaaagac    96060
ttgcactgag ggtgaggtgt taaaaataaa caggtaagta aatgtttaaa catcttgaag   96120
gaaaagtcaa caaatcttgg caagtaaaca gataacaggta aaaaagaatg ggaccaagat    96180
tttgagtttt ggagactggt ggattgaaca gacagggaaa ttgagaggag aatcagatga    96240
tgatgtttta agttgatatt tagacagatt gtgcttgaga tggtaaaagtc aatgtgggtg   96300
ggaatgctta gtagcgagta atcagtgata caagaccaaa gcccaggtca aagacaagtc    96360
acagatacag atcagggctt tttcatctgc tccacagagg tgtaccctag gagctgttgc    96420
aaacagtcca tgtggagggt gtgagtaaga tgtttcccctt gaatttgcca gaattactttt   96480
tttgttgttg ttgttgtttt ttctgagaca gattctcgct ctgttgccca ggctggaagtg   96540
cagtggcgag atcgcgcagc tcactgcaac ctctgcctct cgggttcgag tgattctcct    96600
gcctcagcct cccaagtagc tgggattaca ggttgtggcc accacgccca gctaattttt    96660
tttgtatttt tagtagagat ggggtttcac catgttggcc agactggtct cgaactcctg    96720
gcctcgtgat ctgcctgcct cagcctccaa aagttctggg attacaggtg tgaaccactg    96780
cacccggtcc cttgttaagt ttattttggt gggaagcaaa ggaggtttca gcttttaaaa    96840
agtttgaaaa ttattgctct ggtaataatt aaagatttga gagtaaataat gctttctagc    96900
agaaagaata aaagaagaac agatagcctc aagaaggggga gccaaagaag caggctatat    96960
ctgacacact gggtgttgat aaatgggtat taaaagaatg agagcaatga gcagatagaa    97020
gaggaaatta ggagagtata ataccatgga gaccaagaaa gatagactat caggaaggag    97080
```

38

```
tggtaaaaat aagttactag ttctaagaga gatgttaaga gggaccgggg aaagccttgt    97140
acaaatgagt tagtagcatt ttacattata tacatctaat taagaaacaa tgcgagagtc    97200
tcaccattcc tatagactct tacttgtact tgtctgaaca cgaaaactgg cttttgttta    97260
taaataagct aaaaattatt ttgctccaat ttctcatgaa aataaaaata aaccttcttt    97320
taacattgaa aaaatagttt gaagacagtc actcttcatt ttgtaattcc cacaactatt    97380
attgaatgac tgaaattatc tttattctga agccaaaggg gtgatactga tattttttca    97440
gactactaaa aatatatttt atgaattttt agtgtgcttt atctttttt gttttttttt    97500
ttgagatgga gtttcactcc cgttgctcag gctggagtgc agtggtgcaa tctcagctca    97560
ctgcaacctt cgcctcccag attcaagcaa ttctcctgcc tcggtctccc aagtagctgg    97620
gattacaggc acctgcccce acacccagct aattttttgt attttagta gagacagggt    97680
ttcaccatgt tggtcaggct ggtcttgaac tcctgacctc aggtgatcca ccacacttgg    97740
cctcccaaag tactgcgatt gcaggcatga gccaccatgc ctggcctgag gaatatttt    97800
ctaggttccc cccaccccaa gcatttattc tgcaatttta gtttgttcc taaagcaagc    97860
tatccacagt agaagttctc agagaatgat ctccctcttt taatttaact ttttggcaca    97920
gtattttgag aattataaat aatattagaa tgtttctgg ctggttgtgg tggctcatgc    97980
ctgtaatcct ggctacttgg gaggctgagg caggagaatc acttgaacat gggaggcaga    98040
ggttgcagtg agccgagagtc atgccacatgc actccagcct gggtgacaga gcaagactct    98100
gtctgggaaa aaaaaaaaa aaaaaagag tgttttcttt cctattttcc accacttgat    98160
taagttactt ttcctcttaa gtattttttg ctgagtatgc tgacttaaga gtaatgttac    98220
aaaatttaat tttttaaagtt ctctgaaagc ccctttatga gagttttagg ctatcaaatt    98280
gtgtttaatt cttaacaatt ttttgaaaaa ttatagctc aatatccgta cattccccac    98340
aaaaaagcac taaaaatcat gccttgctgg aggctgcagg accaagtcat gttgcaatca    98400
atgccatttc tgccaacatg gactccttt caagtagcag gacagccaca cttaagaagc    98460
agccaagcca catggaggcc gctcattttg gtgacctggg taagtaacta tcatttttta    98520
ttaacttgta ttagaaggat ttgagtacaa tatgtgaaac ttctgtcata ggatacagaa    98580
ctatataatt ggaaagtgct ttggaaaaaa tgtatttaaa ataacagcta caagtataat    98640
gggtagctgt gttgtgttcc tgtaaatata gaatataaag catgcccagt agaaaaacaa    98700
gcatttccag aagaaaata tctgatcact aaatataaat atatgaaaaa gatgtctcac    98760
tttattactg agggaagtgc aaattaaaat aatcagttaa tgttctccta acacattagc    98820
atattttta aagtttgaca atttgaatgt cagtgaagat gcagggaaat acccctccta    98880
tttagtgata atataatctg gtgaagactc tttggaaagc aattttgaaa tcagtataaa    98940
atatgcatgt catttaggcc actctttcta agacctagcc ctcagatatg ctcattcata    99000
tgtgcaggtg tgtatgtgtg tgtgtgtgtg tgtgtgtgtg tgtatatgta tgtatgtatg    99060
tatgtatgta tgtatgttga aggctattca ttatagtatt gtttgtgata gcaaaaaatt    99120
atggacaaca tataaatatc tgttataggg aaataaccaa attgtggtat acgcatgctc    99180
tggagtataa tatagccatt tgtttctatt tatttatttt cttgagacag ggtttttactc    99240
tgttgcccag gctggagtgc agtggtatga tcatggttca ctgcagccatt cacctcctgg    99300
gcacaagcca ttctctcgcc tcagcctcca gagttactag gactgcaggc atgtctcacc    99360
acacccagat aatttttaa ttttttgtag agacaggttc tcactatgtt gcctaagctg    99420
gtctcaaact cctggcctca agcaattctc ccacacaggc ctcccaaagt gctgggatta    99480
ccaacgtgaa ccaccacacc tggttcagtg tagccattta gaaatctaaa aaagacgtgg    99540
gaaaatgtct aaggcatgtt taaatgtgag aaaagcaagt cacagtatgc atggtaaaat    99600
ccgttatatt aaaataagtt cttccaaaac aaaaacatat gcaggagacc tttatttgt    99720
cagtatttct taccaaatt tctgcactta gaaaatttgca tgtcatgttg tcataagttg    99780
aaaaaaaagat ccatgaacca atggacttct aataaaatca gtcctgcttt tgacatctct    99840
ctctactttt gtgtatattc aaaccagagt gtcaatggat tgtgtgggca cacttagcaa    99900
taatacatag cagacaaaat gcataagact cagagagtaa aattgtaagt tttgctgaat    99960
cactcataaa ttgctgatga gaatttaaaa tcggtcgaga gctctggaaa acaggcagtt   100020
tctttcttttc ttttttttt tcttttgaag acagggtctc actctggtgc gcaggctgga   100080
gtacagtggc gtgattacaa ctcactgcag acagggtctc actctgttgc ccaggctgga   100140
cctcagtctc ctgagtagct ggactatag gcatgcacca ccacgcctgg ctaatttttg   100200
tattttttt ttttttttt gtagagacgg ggtttcgcca tgttccccag gctggtctca   100260
aactcctgga atcaaggcat ccacttgggc aggcctccca aagtgctggg attacgggcg   100320
tgagctactg tgcctggcct aggcagtttg tttgtttgtt tgtttgtttg tttatttatt   100380
tgtagacaga gtctcacagg ctggagtgca gtggcccaat ttttggtca ctgcaacctc   100440
cgcctcccag gttcaagtca ttctcctgcc tcagcctcct gagtagctgg gattacaggt   100500
gcctgccacc atgcctggct gattttgta tatttagtag atatggggtt tcaccatgtt   100560
ggtcaggctg gtcttgaact cctgacctca ggtgatcgac ccgcctcggc ctcccaaagt   100620
gctgggatta caggcgtgag ccgtcatacc tggctgttgg ttcttatga cgtgaaacat   100680
gcaattacca tatgacctag cagttgcact ctgtatttat cccagataaa tgaaaactta   100740
ccttccaata aaaacctgtg cacaaatgtt catagcagct taatattgaa aaactggatg   100800
ttcttcagca ggtgaatgaa ctggttcatt cataccatgg aataccattc agcaataaaa   100860
aggaacaaac tgttgataca tttaaccacc tggatgaata tcaaggggaat tatgctgtca   100920
gacaaaaacc agtccctaaa gactacatat agtatgattc cgtttggata atattcttga   100980
aatagagaaa ttaagagaaa tgaaaagatt agtgtttgcc agatgttaga gacagggagg   101040
tgagaggggt aagtgggtgt agttatgaaga gtgcaacatg agggatcttt gtgatgttga   101100
```

                                         39

```
agttgtatct tggcagtgga tgcagaaatc tcaatgtgat aaaattacaa agaactaaaa   101160
acaagaatga gtatagataa aactggggaa atctgaacaa gttagagtgt tgtatcactg   101220
tcagtatctt agagtgatat tgtactatag ctttgcaaga tgttaccatg ggagaaacta   101280
aagtgtacaa gggatctcta ggtattatta tttttttaga gatggggttt cactatgttc   101340
cccaggccgg tcttgaactc ctgggctcta gtgatccgcc tgccccagcc tcctaaagta   101400
ctggaattac aggcgtgagc gaccatgcct ggcccttca gtattgtatc ttagaacttc   101460
atgtgaatct agcattatct catagaattt aattaaaaga aattgtaaac ctcacagaag   101520
atcagaattt cctcaagttt gtgatgttga caaagatgaa ctagttgaca ctggagtaa   101580
gactgaggat gaagacacga cgtgcttcaa aaaaatgatt tgaatatcaa tggattaaga   101640
agaactcttt tgacaaattg atgaaaccct cagtcagttt tataagaatg cccatctta   101700
tgatcatgct atgaaagcca attttaaaa aaatttttg tcttttctaa caattagctt   101760
gtggttataa tttaaattta gttaaatata agataaatga tttttatta agtttagttt   101820
cattttcaa ggtacgatct caaagctact ctttaaccta ctatgaatga ataatgctga   101880
gttcataaca tctttgtaga tatatccaca attttccctc aggataagtg cctacaagtg   101940
gaattactgg actgaaaata atgcagtttg ctaagacttt gctatctgtt cctgaatgct   102000
cctccaaaaa ggttttgcca gtttacatcc tcatgaccag cgaatgagag tgttgcctat   102060
tttcctgtgc ccttgttact gcttaataat ttttgaaaaa aatctaattt gacagacaaa   102120
aatgcatttt atgttaattt gcttttctgg gattttaat gaggttgagt atagttttta   102180
atattttat tggccccttt ggaactagta tcataagttt tttttcttaa gaatttatgt   102240
agtctgggct gggcgacgtg gctcacgcct gcaatcccag cactttggga ggccgaggtg   102300
ggtggattgc cgaaggtcag gagtttgaga ccatcctgac caacatggtg aaaccgaatc   102360
tctactaaaa gtacaaaaac tagctcagcg tggtggcggg tgcctgtaat cccagctact   102420
taggaggctg agtcaagaga atcgcttgaa cccgggaggt ggaggttgt tgcattgagc   102480
cgagatcgcg ccattgctct ccagcctagg caacaagagt gaaaagtctc aaaaaaaaaa   102540
aaaaaaaaaa aaataagaat ttacatggtc tgaattgcca ttaaaagaa taaaaatgaa   102600
attgagtaac aaataacttt ttaataattt aggcaagttt tggacgattg tactttgttt   102660
agaaaccaaa agcatagtat ttgtagtttt tttatttact ttagttgcta ggaagtaaac   102720
tttattcaag gtctctggta ccagttgttg ctaaaagtga ttgactaatc tgtcaatctg   102780
aaattattty ttgctgaact gctaattctt ttgcttctat cttttaggca gatcttgtct   102840
ggactaccag actcaagaga ccaaatcaag ccttttctaag acccttgaac aagtcttgca   102900
cgacactatt gtcctccctt acttcattca attcatggaa cttcggcgaa tggagcattt   102960
ggtgaaattt tggttagagg ctgaaagttt tcattcaaca acttggtcgc gaataagagc   103020
acacagtcta aacacagtga agacgagctc actggctggg cctgtctctc catctaaaaa   103080
gcatgaaact acagcgtctt ttttaactga ttctcttgat aagagattgg aggattctgg   103140
ctcagcacag ttgtttatga ctcattcaga aggaattgac ctgaataata gaactaacag   103200
cactcagaat cacttgctgc ttcccagga atgtgacagt gccattctc tccgtcttga   103260
aatggccaga gcaggaacctc accaagtttc catggaaacc caagaatctt cctctacact   103320
tacagtagcc agtagaaata gtcccgcttc tccactaaaa gaattgtcag gaaaactaat   103380
gaaaagtgag tatgtgattt tcttgtgtgt acatatgtgt ctcactttct ttttttaatt   103440
tactaagcag aacttcagat gaggaataaa atgattggaa tatttttt ctcctctaac   103500
tactgtaaa tttgggagaa tttggagagt gtagtgagt cagatcagtg taggttgac   103560
gagcaggagt gactggacct tctaagaagt gtgttatcag aattagtaaa tgaagggtca   103620
aatgtcctac ttttccctc cactgatttt gacatcaaac cattatccac atagccttat   103680
ttcctccctc ggcttaattt ttattaaata tttactgcac tttgcagata aaatttttaa   103740
aaaattttta aaaattgcca ataagtgaca tttattaagt tcagtgctta gtgtatattt   103800
ggattttatt tattagtcac aagacctttg tgcagtgagt agccatgatt atctttttt   103860
ttttgagatg gagtcttgct ctgtcgccca ggctggagtg caatggcgcg gtctcggctc   103920
actgcaacct ccggttcat cgccattctcc tgcctcagcc tcccaaatag ctgggactac   103980
aggcgcctgc caccacaccc ggctaattt tttgtatttt tagtagagac ggggtttcac   104040
catgttcgcc aggatggtct cgatctcctg actttgtgat ccgcctgcct cggcctccca   104100
aagtgctggg attacaggca tgagccaccg cgcccggact gattatctta tttacacatg   104160
agaaaaccag ggcttagaaa ggttaggtaa cttcctctag gttgtacagt aaatgtggac   104220
ctagaagcat tttgacaaga gcacctgttt tttttctttc tctattagtt tagaaattat   104280
atactcttaa ttatcacctg ggattttgat tagacagcct tcatgttctt tttcatctta   104340
aatgttcttt gtgtcttaaa gggctaagtg atttcttcag atcttttagt tcactcattc   104400
tcagtgaact aaaatgaggt ctaatctgct actgaatcaa gttttcagca tgttatttcc   104460
ttcctccctc cctccctcct tccttccctc aaccaggctc ccgaggagct gggattacag   104520
gcgcccgcca ccacctcctgg ctaattttta tattttagta gagacggggt ttcaccatgt   104580
tggtcaggct gatcttgaac tcctgaccctc aagtgacccca cctgcctcgg cctcccaaag   104640
tgctgggatt acaggcatga atacacacac ctgacggcat gttattttca tcgcaaagtt   104700
actgtaagct gggagaagtg gcacacactt gtactcccag ctactcagga agcttaaggt   104760
gagaagattg cttgagccca ggagttttga gaccaacctg ggcaacacag caagacccca   104820
gctcaaacaa agaaaaaaag ttattgaatt tttatttct atggatcatt ttttgtagtt   104880
tcttattcct ttcacccttc attcccactt ttgatcccat cttttattta tttagtttta   104940
ttaaatgtat atttgtctga taattctgct atctacagtt ttttgtggac ctgactcagc   105000
attctttgt tcttcggat tcagactgtt ggtggcttgt gattttagtg atttttggcc   105060
gtgaacatgt ttcttggact tttgtctgtg ggaatctctct gtgtactctg tataaattaa   105120
```

40

```
gttacttcag gtgttttgca ttttcttttg ccatgcacct ggggcctggg tcactaccct 105180
tctggtacca cttaaaactg aattttgtc ttgggtgctc gtactgatcc tgtatgagta 105240
caggtttata cttactgtag aaatatggtg tttgattatg gggtattgtc ccagatggtg 105300
ctggagtatt aatatgctct ctgttaaact taatgtgttg tccctgtaaa actccaaaat 105360
tctgaattcc agaatactac tggccccaaa tgtttaagat aagggcactg cctgtatttg 105420
tttctgcctc ccactatttt ccttagttta acacaaactc accttttaa aaaacatttt 105480
gagagaattc agtattggga agagtttcta acctgtttct ggaaatggaa gtccaaagtc 105540
tgtttctgta attgtttttt ttttgagatg gagtctcact ctgtcaccca ggctggagtg 105600
caatgacgta ctctcagctc actgcaacct ccacctcccg ggttcaagcg attctcctgc 105660
ctcagcccca tgagtagctg ggattacagg tgcccaccac catgcctggc tgattttttgt 105720
attttttagaa gagatggggt ttcgccatgt tggccaggct ggtcttgaac tcctgacctt 105780
gtgatctgcc cacctcagcc tcccaaagtg ctaggattat gtttctgtaa tgtaataca 105840
tttattgttt ttagaaactg tctttgcttt agtggtaaat tccaataaaa atagaaatag 105900
cagtggagtt attaaaagag cattagttac attttccct ttttcattat cttcaaatat 105960
tatatatagt aagtttgacc tttttaaaat gtatacttgt atcagttta acacatacat 106020
agattcctgt aactgtcacc actataaggg taaagaacag ttagttcctt caccttttgaa 106080
gtcaagcccc acctcatcc caacacttgg caaccgctga tctttctccg tctcaatagc 106140
tttgcctttt ctcttttttt ttcttatttt tttttttgag acagcgtctt gctctgtcgc 106200
ccgagctgga gtgcagtgag gcaatctcgg ctcactgcaa cctccgcctc cctggttcaa 106260
gcagttctcc tgcctcagcc tccctagtag ctgggattat aggcacgcac caccacaccc 106320
ggctgatttt tttgtattct tagtagaaat ggggtttcac catgttggcc aggctggtct 106380
caaactcctg acctcaagtg atccacctgc ctcggcctcc caaagtgctg ggattacagg 106440
cgtgagccac tgtgcccaat caggactttt tttttttaaa tttacattca acttgtcatt 106500
tttttcttgt atggattgtg ccttcagagt cacacctaag agccctttgc ctaagcaaag 106560
gtcatgaaga ttttctcata tgtttccttt taaaagtatt gtggtttggcc aggtgccatg 106620
gcttatgcct gtaatctcag cactttgaga agctgaggtg ggcagattac gaggtcagga 106680
gatcgagacc atcctggctca atgcggtgaa accccatctc tactaaaaat acaaaaaaaa 106740
aaaaaaatta gccgggcgtg gtggcgggca cctgtagtcc cagctacttg agaggttgag 106800
gcaggagaat agtgtgaacc cgggaggtgg agcttgaagt gagccgacta gccgccactg 106860
cactccagcc tgggcaacaa agtgagactc catctcaaaa aaaaaaaaaa agtattatgg 106920
ttttacactt tacgtttaga tatatatctt ttttgagtta atgtcgtata agtatgaggg 106980
ttacgtcaga tttttttgttt tttgtttatt ttacatatg gatgtctagt tgttctaata 107040
ccatttgttg aaaagacaac ctttactcca ttgaattgcc tttgtacttt tgccatattt 107100
gtctaggcct gtttttggac tccttttctct gtttcatgat gtgtgtgtct attcctttgt 107160
taataccaca tggtcttaat tactgtatag taagtcttaa aattgggtaa tgctggcctt 107220
ataaaacgaa ttgggaagtt tttattttta ctcttatttc catttttctag aagagattgt 107280
gtagaattgg tgtcatttct tcctttagata tttggttgaa ttgggaagtg atgccatctg 107340
ggcctagggt ttgtttttt gtgtgtgaga cagagtctca cttctgtcac ccaggttgga 107400
gtgcagtggt gagatcttgg cttactgcaa cctctgcctc ccaggttcaa gttatcctcc 107460
tgcctcagcc tcccaaatag ctgggattac aagcgtgtgc caccatgccc gactaatttt 107520
tgtatttta atgcagacag ggtttcacca tgttagccaa gctggtctcg aacttgtgac 107580
ctcaagtgat tagcccacct tggcctccca aagtgttagg attatagatg tgagccaccg 107640
tgcctggcag gggcctaggg ttttctttt cagagtattt taaactatga attcagatta 107700
tttaatagat ataggactat ttaagttatc tgtttcttct tgagtgaatt tttacatgag 107760
tttatgcctt ttgagtaatt aattgtattg aattgtcaaa ttatgagcg tgtaattatt 107820
tatagcattt cgggtttgta gtggtatccc tctttattcc ctggtgttgg caattgtgtc 107880
tgtttttct ttgtcagatt gatagggat ttctgctctt tattattct catttttct 107940
ttttgattt tctgtcgttc tttgggttta tttctactct tttttttttt ccaagttgct 108000
taaagtagaa acttagttt ctggtttgag accttttctt tctaagataa gcatttaata 108060
ctgtaaattt ccttctaacc actgctttag ttacaccccc acaaattctg gtattttgaa 108120
ctgagcacaa atgaaatgtt ctaattttccc ttgaatctta ttcttttacc aatgaattat 108180
ttagaaatat gttatttagt ttgcaagcaa ttggagactt ttttcctgtt atttttctac 108240
catttatttc tcatttcatt atattatggt cagagaatat atttttgaatg atttcattta 108300
ttaattttta aaaataacat taaaaaattt ttaaaatgt gaatatacca catacagtat 108360
aaagattgta cattctgttt ttggacagtt ttctataaat gtcaagttga tttagtttggt 108420
taatgatggt gttcagtttt tctttattct tgctgatact ttgtatgcag ttatatcact 108480
ttattactca gaagagtgtt gaactttcca actacaattt tttttttccaa ttttactttta 108540
agctctatct ggtttttgctt catgtatttt gaggctctgt tgttaggtcgt gtacacattc 108600
aggatgatat cttctggggtg aattgcctgt tttatcatta tgtaattccc tctttatggt 108660
aatttctctt gttctaagat cagaaatatc tgttgtccaa tttatataga cactgcagct 108720
ttcatttgat tagtgcttgc atggcatatc tttttccatt tttttacttt tgatctacct 108780
ttataattct atttaaaggg ggcttctttgt aggcagcata tagttgggta gtgttatttta 108840
tttattaatt tattttattta tttatttatt tattgagaca gagtttttgct cttgttcccc 108900
aagctggagt gcagtggtgc aatcctggct taccacaacc tccacctct gggttgcagt 108960
gattcctg cctcagcctc ccaagtagct gggattacag gcacgcgcac catgcctggc 109080
tgatttttg tattttttagt agaaacggat tttcaccatg ttagccaggc tcgtcttgaa 109140
```

41

```
ctcctgacct caggtgatcc acctgctttg gcctcccaaa gtgctgggat tacaggcgtg  109200
agccactgca cccggctgag tcatgttatt tttaatcttt tctcacaata cagggttttt  109260
gttggtcactg ttaattattt taatataaat tttagtataa tatttacat taaatgtaac  109320
tgttgcactg gggtatttat aatgtgtaaa tataattatt ggtattaata taattatatt  109380
actcataata atattaatat ctttggattt agattaccag tttagtatat gtttttctgt  109440
ttctccctct ttgatttccc cttttttgct tttttttttt ttttaattct tattttttttt  109500
tagtatttgt tgatcattct tgggtgtttc ttggagaggg ggattggca gggtcatagg  109560
acaatagttg agggaagtc agcagataaa tgtgaacca aggtctctgg ttttcctaga  109620
cagaggaccc tgcggccttc tgcagtgttt gtgtccctgg tgacttgaga ttagggagtg  109680
gtgatgactc ttaacgagca tgctgccttc aagcatctgt ttaacaaagc acatcttgca  109740
ccaccttaa tccatttaac cctgagtggt aatagcacat gtttcagaga gcaggggtt  109800
ggggtaagg ttatagatta acagcatccc aaggcagaag aattttttctt agtacagaac  109860
aaaatggagt ctcccatgtc tacttctttc tacacagaca cagtaacaat ctgatctctct  109920
tttcttttcc ccacatttcc ccctttttcta ttcgacaaaa ctgccatgtt catcatggcc  109980
cgttctcaat gagctgttgg gtacacctcc cagacggggt ggcagctgag cagaggggct  110040
cctcacttcc cagatgggag agccgggcag aggcgccccc cacctcccag acggggcagt  110100
ggccggcgg aggcgccccc cacctcctcc ccggatgggg cggctggccg ccccggggct  110160
gaccccccac ctccctcccg gacggggcgg gacggggcgg ggggctgac ccccgacctc  110220
cctcccagat ggggcggctg gccgggcggg ggctgccccc cacctccctc ccggacggggg  110280
cggctgccgg gctgaggggc tcctcacttc gcagaccggg cggctgccgg gcggagggcc  110340
tcctcacttc tcagacgggg cggccgggca gagacgctcc tcacctccca gatggggtgg  110400
cggtcgggca gagacactcc tcagttccca gacggggtcg cggccgggca gaggcgctcc  110460
tcccatccca gacggggcgg cggggcagag gtggtcccca catctcagac gatgggctgc  110520
cgggcagaga cactcctcac ttcctagacg ggatggcagc cgggaagagg tgctcctcac  110580
ttcccagacg gggcggcccg tcagagggc tcctcacatc ccagacgatg ggcggctagg  110640
cagagacgct cctcacttcc cggacggggt ggcggccggg cagaggctgc aatctcggca  110700
ctttgggagg ccaaggcagg cggctgggaa gtggaggttg tagggagctg agatcacgcc  110760
actgcactcc agcctgggca acattgagca ttgagtgagc gagactccgt ctgcaatcct  110820
ggcacctcgg gaggccgagg caggcagatc actcgcggtc aggagctgga gaccagcccg  110880
gccaacacag cgaaaccccg tctccaccaa aaaatgcaaa aaccagtcag gtgtggcggc  110940
gtgcgcctgc aatcccaggc actctgcagg ctgaggcagg agaatcaggc agggaggttg  111000
cagtgagccg agatggcggc agtacagtcc agcctcggct ttcacaactt tggtggcatc  111060
agagggagac cgggggagag gagagggaga cgagggagag ccccttttttt gctttcttt  111120
ggattatttg aattttttcct taaatttatt tatcttactt atttatttat tttttttgagt  111180
gattctcctg ccacagctcc caagtagctg ggactgcaag catgtgccac tacacccagc  111240
taattttttt gtatttttag tagagacgga gtttcaccat attggccagg ctggtcttga  111300
actcttgacc tcaagtgatc cacctgcctc ggcctcccaa agtgctggga ttacaggcgt  111360
gagccaccat gccctgcctt tttctagaat ttatatattg agttcttgat tgtatctttt  111420
tatgtaggct ttttagtggc ttctctagga attacaatat acatacttttt cacagtgtac  111480
tcacatttaa tatttttgtaa cttcaagtgg aatgtagaaa acttaaccac cataaaaata  111540
gaactaggga tgaggttaaa aaagagagag aaaagaaatg taataaagat ttaataaagt  111600
cgttttttttt tttttttctc tttttttttttt gagacagagt ctctctttct gttaccaggc  111660
tggagtgcag tggcgtgatc ttggctcact gcaacctccg cctcctgggt tcaagtgtt  111720
ctcctgcctc agcctactga gtagctggga ttacaggtgc gcgccaccat gcccagctaa  111780
tttttgtatt tttagtagag acggtttcac tgtgttggcc aggatggtct cgatttcttg  111840
accttgtgat tcgctctcct cagcctccca aagtgctgga ttacaggcg tgagccaccg  111900
cgcccggcta agtctttaaa tatttttttg acattgcact tttttctttt tccttctaag  111960
attttagtaa cccaaatgtt agttttgta ttgtttggca ggttcctgag gcttctcctta  112020
cttctttaaa tttttttttc ctgttgttca tctgttattg agtctttgta gtgaatttta  112080
aatcactgg ttctttcccg ttatttccat tctcgttattg agtctttgta gtgaatttta  112140
aatttttgttt attatgtttt ttagttctaa aatttttcttt ttttttgtgtat gtcttatact  112200
ttgctcctga aactcttatt tgtttcagga gtgatcttat ttcttagagc atggtttttag  112260
ttgtgagata atgggatttt ctcgttctttt tatcatccca gcatatgtgt cctcttgatt gtcttttctc  112320
gacagtttga cttaacgttac atgattctga atcttgttta aatcctgtgg aaaatattga  112440
agttttttgct tcaacaagca gttgacctag ttagtttcag tccacaaatt ctaagcagca  112500
tctctgtcgg tctggttcca tcatcagttc agttttgtat cttatctgct tatgtgcctt  112560
tctgtgtcca gtctgggacc tggccaatgg tcaggtccca aagcctttgt acacttttag  112620
aagcagggcc atgcacaccc agctcacgag tggcccgggg agtgcacata caactcgacg  112680
ttttcatggg ctccttcttt tctgtgatgt ccctgacacg ttctgccttc taagaacctc  112740
cctttatccc tttcctgttg tctggctaga aagtcagggc tttagattcc ctatacttca  112800
gcacacttct cgtagctatg tcaacctctg tggccacgac ttcttcttct tgggactgca  112860
gtttctcttg tcagaaagta ggattcttgg agctgctgtc attgctgctg tggctgctct  112920
gatgctgcct gggagtcgaa ggagagaaag gaacaaaaca aacaacccaa atagagggc  112980
tccactctct ttgatccgtg agagcccct ttcctgttcc tcagaccaga aatagagggc  113040
ctgtcttgga acttctttctt tgtgcatctg gtgtgcagtt tcagcttttg agtccaggcc  113100
aggaggtgct ggacaaactt gtcaggagta cggaggtact gcaagttctg attactttc  113160
```

42

```
tcagtccacc tgcttccaag tccttggatg catttgtcca ttgttttgag ttgcattcca  113220
tgggagagac agaagagtgt gcttatttca tcttgacata cttattagga tttcatatca  113280
aatcaacgga tgatattctc tatattaatt tgctgttttc cctttagcaa gcacattagg  113340
aaaataacac tttaacaccc gccttggtg gtttctgtca taattattaa tacttgactt  113400
ttttttttttt tttgagacgg agtctcactc tgtccttga ggcattgtcc ccataaactt  113460
ttggtaaagc atcaataatt ttatctttca tccacacaag cttcaccata aatttgatgt  113520
ttattcttcc attttagcag aattcatgtt gctccaatag gggctgtctt caaactgatg  113580
ttttctccctt cttagtgcct cagagtagat ccgtttagaa tacgttataa caggttaata  113640
tgagtttatt ttggtgtaaa agtactttga aattcatgca tagtttttc atcataatcga  113700
ttttccatag ctttgaacac ccccatgtaa ctctcctctt ccacaaacca aacaatgaaa  113760
aagcaccttt gtgatggaag tttatttgc aataggaact cacagtgatc taagccccgc  113820
tattcatgaa tataattcat tactggagtc caagttgtt ttggtttct gaagttctct  113880
tcttccctg caggtataga acaagatgca gtgaatactt ttaccaaata tatatctcca  113940
gatgctgcta aaccaatacc aattacagaa gcaatgagaa atgacatcat aggtaagcag  114000
tgcttgaaac tatggcaaaa aaaaaatgac aaaaatgca cagaactgac aatttttgtt  114060
attgactaag ataattttt cttaacatgg aatttagcag ttcccttcct aatttgtttt  114120
ctgagtattt tttatatcgg attatagctc acttttaaaag tttctcggct gcattcggtg  114180
cgagggtctt tgcctgggcc agatgggctg cagtgtagcg ggtgctcagg cctgcccgct  114240
gctgagcagc cgggccggcg ggcggctacg ctaaccggca cagcaccg gggctgcccgg  114300
ccggcagccc cgcaccagtg cacgaagtgg gcgggacaga aacttctggg gttggaagtc  114360
cagtgaggct aaaagccgt accaaagtct ctaggcatca gggctgcaac ccaagagtct  114420
cacgaccagt gggcaactgg atggccagac aggtgtctca gtggtggcct ctccgtctca  114480
gggcttcatc ccacttctca gtgggcctga cgtccctggg cacccggat gtctacctgc  114540
attagccaga gccatcacat ggcctgtgac ttgcctttt ttgccagttg attgtgccac  114600
acacagtgtc atttctgtgt catttggcac agctggaggt gcaaggagga gggcagccct  114660
atgtccagtc ccagtttcac gtaactttat tcttctgaat aaagacaatt tgctaaccctt  114720
aaaaaaaaaa aaaaaaaaaa agtttttctt atatgttgga cccaaattct taggcttttaa  114780
cctgaataac aatgacagca agatcaataa atagtacaca tttattaaac actcactgtg  114840
tcccagacaa tattccaagc acttttttatg gatagactcc ttttaacttc taaagaactt  114900
tgtgggataa atacagttat tttatgatg aagaaactga agcacagaga agttaagtgc  114960
tttgtccagg gtaacagctc agatatggca gagtcaggat ttgaaactag accctcacat  115020
accttaactg ctgtgctgtg gcagtgtttt tcatatctga ggttgggacc agccttctct  115080
tatgccctca cccctgcca aaaaaaaaa aaaaaaaaa aaatatatat atatatatat  115140
atatatatat atatatatat aatatatata tatataaaat atatatatat ataaaatata  115200
tgtattagta tatatgcata tatagtatat attatatatt agtatatata ctaatatata  115260
atatacatat tagtgtgtgt atatatatat atactagaat aaaaaaatca aagtatctca  115320
gagtagtaag gacaaacatt tcagaaaaat gttttcatta tatatacatg tatgtatgtg  115380
tatgctgatt caacaaatat atttcttata ggtatagca aaatagtttg aaagcttttta  115440
ctgtgtttta tcaggagac cttaggtgaa cgtatattca cagataaaag aggttattta  115500
ttcattcaat aaatattaca ttctcataag tcctaatatt atgtattttt attcttcaaa  115560
aaagttagta tttgtgattt atgaaataag acatgttctt gcactttag cagatctgtcc  115620
ccgatgttgg gcttctttaa tccttagtgt gggtgctttg cactcactca ctgctggga  115680
cagcaagacc cctgttagtc tcagctgtgt ttcttaaatt ggcccactgt acctccagt  115740
tagctattct gggtccatg tcatgttggc tccattttcc ttttctttct cccacacaga  115800
tacctataac ggctataaca taggcctggt ggctgttgat ggctatccc tatctgcttg  115860
tatttaaggg gtactgtttc actgagtttt gctgacagat gttgtcatga gatttgaggt  115920
gctttttccta gttgctctat ttttatgtgg gaattttgcta tcatcatcat ccctagacca  115980
ttggttggct agtgattttt ttttgagggg agtcgtacc agttaatagc ctgactgggcg  116040
gcttttccta gtaatacaac agggatgttc tgatctgacc aagtttgcct gtttgaagaa  116100
tgtggataaa aaggaagcag tttcaagtca aataaacacc ttaaaatgaa accacactgc  116160
aactctcttt cttttactta agttaatca aattaatgat gatgtaatcc catgaaggaa  116220
aagtcttctg aaggatcaag ttgataacat tttgtgatca aagaatttga gaaaacctct  116280
atcccagtgt ctatcattat atatttttagg atgttaatta cctgtgtggc tttaggcaag  116340
tcattttctc tccttgagcc ccattcttaa tcctgtccaa attatttgtc tcctcttgca  116400
gttggactat tttaatatag ctgtccttca agtgagtttt gttcaaagga gccttcactt  116460
tagctcttac tgtgtaccca ctttgcatag tcttgttttta aatgtaatcc ttgggggtttg  116520
ggtgttgcta actaattact gttttatgt gaggatttag agtgatccag aatctatact  116580
tgcactacct ccttcatctt ccacaaatgt ttgaagtggt agaattttta aaaacttttga  116640
aggtacagct gacagaattt gctgatggtt tggaagtgga tggtatgaga gggaaaaaaa  116700
ggaataaagc atgactgcat ttttttgttt ttttgttgtt tgtttttgtt tggtatttag acggagtctc  116760
actctcgcca ggctggagtg cagtggcgtg atcttggctc acggcaacct ccgcctcctg  116820
ggttcaagcg attccctgc ctcagcctcc caagtagctg ggactacagg cgcctcgcca  116880
cacgcctggc taattttttt ttttgtattt tagtagaaac ggggttttca cgtgttggcc  116940
aggatgggtct ccatctcctg acctcatgat ctactcacct tggcctccca aagtgctgag  117000
gttacaggca tatatataag catataaagt gtgttatagc atacaaacag gtatatatat  117060
aaacatgcag tccacacagc tgataggaat gaggcagtag tgaaggagaa gttgatgtag  117120
gagagggac agttgttaca ggaaagaagt ctggaggcag aagggatgaa ttccagtgct  117180
```

43

```
cacatagaag attgcttaga tgggagcaag gacaatttat ctagagtcac aggaaagaat 117240
gcagtacacg ggtagagatg caggtgagtt gaaagatgtg agagatgatg gaaataattt 117300
tctgattgct tctatattct caaggaagca ggaagcaaag tcctcagcaa agagaataga 117360
agaggtgtta aatatttgag aaaggagatg tactgtagaa aaaaaaaaaa ctcagtttct 117420
ccttctgaac tctcacaaaa cagaaccctt ccatgactct agttgtgtgg ggttttttcc 117480
ctgtcagcta ccaattctgc agatgattgt tcagtgaaca ccaactgggt gtcctctaag 117540
tcagttcagt tctcacactg tttacctgga gatagcatca gatcccacag attgaggacct 117600
ctgtcccaca agactgcctc cacttcagat gccagtctca agtacaagtt gtgggcctgtg 117660
cttctgactg accttctata aattggagtt cccacagtcc cctccttggg ttaataaat 117720
ttgctagagc agctctcaga actcagggaa atgcttttaca tatatttacc cattctatat 117780
aaaggatatt acaaaggata cagattgaac aggcagatgg aagagatgca tgggcaaggt 117840
atgggagagg ggcacgaagc ttccatgcac tctccaggtc atgccaccct ccaagaacct 117900
ctacagattt agctattcag aagcccccct ccccattctg tcctttgggg tttttgtggg 117960
agacttcatt atataggcat gattgatcat tggctattgg tgatcagctc aaccttcagc 118020
ccctcatcc cgggaggttg gtgggtaggg ctgaaagtcc caaacgtgta attctgcctt 118080
ggtcttctg gtgattagcc ctcatcctaa agctcttag aggccacagc cacaagtcat 118140
ctcattagcc ttcaaaagaa tccagagatt ccatgaattt taggcgctgt atgctcagtcc 118200
actggctaaa ggccagttgc aatgtctcag gcctgtaatc ccagcacttt gggaggctga 118260
ggcaggagga tcgtttgagg ccatgagatc aaaaccagcc tggtcaacat ggtgagaccc 118320
ccttacaaaa aatttaaaaa ttggccaggc gtaatagctc ttgtctgtag tctcagctac 118380
tcagaaggct gaggatcagct gagccctgga gttgaaggca gcagtgagcc atgatcgtgc 118440
aaaactgggc aaagactaaa taacatattt cacagtatca cagatttgta ttgtctagga 118500
aagtgaatgt aaacagacca ggacactagt atgatcccct ggtttcatga aggtcccact 118560
aaagtcatga acacaaagtg agactaggca tcatgttata tggtttttcc agccatgttt 118620
aacagctagc taaatagctа attgttcgc tgcagtttat tttagcagtt cttattttа 118740
gcacatttca tgttttaaaa tttctaccaa taacatttta ataaactttt ttacagataa 118800
cttcacaaat ccataatttt ttaagttaca atcccagaaa tagaattgct cattgaaagg 118860
gtatgttcat ttttaaagtt atgctagaaa ctgccaaatt gccttcagaa aaaggtgttt 118920
gtatcccсac taacactagt gttagttttc ttgtgccctt gctcaagtat acatattatt 118980
aaaaaacaatg ttgggccagt ttactagata aaaggtgtag tgcctcctta ttctaatcta 119040
tttgattact agtgagtatg tatgtctttt cacgttggtc attttatgtt tgttcctttg 119100
tggattgtca tgtcctttgc tcatttttct tttggaacat ttcttagtag ttttataagag 119160
ctcttggtat tttaatgata gtaacctttt aactgtcatg catgctgcaa atcttttttc 119220
tgtttgtttg ccttttgtatt ttgtttttgg aggggtttcta tgtataggaa ttaaattta 119280
tgttgtaaa tcttttgatt tctgcttttg catatgtact tcaaaagact tctattttta 119340
agatcaagtg ttaactgtat tttcttttag ttctatttaa aacctcttaa tttatatgcc 119400
tgtgctgtta actcccaagt tgattcacaa gtgtgtatac atagtttgaa tttagtggca 119460
atttaattat ttacaacttc ttttgcagca aggatttgtg gagagatgg acaggtggat 119520
cccaactgtt tcgttttggc acagtccata gtctttagtg caatggagca aggataagtt 119580
agttcatatt ttacattgc gcatcctagg gaatttgggt tcattgttag gaatgggctt 119640
cactcagcta aaaacaaagt aattttgaga atttaaatat tttggatatt tacaagatca 119700
tataaagcat actctatctt ggttaacagt ttctttaaaa tataaattat ttgtttttag 119760
aaaattttca ttttcatttt caatgttaat atttcctaag ttaaaataat ttgttttttag 119820
ttctgaaata attttggggag tgattgagtc tgtagtgatt atgactatta gaattggtttt 119880
atttattttaa ataatgcatg tcttcagatg gctctcctaa tttgttagtt aggctttaag 119940
ctaaatggat gctatataac taaatccaca tagatttgtt gaaatggctc cagaggtttt 120000
ttagatttat tactgctatg tgcccttaaa aaaaatctat tcattctttc acttaacatt 120060
tatcagaagga gtgctctgtg taagacgtgg ttaggcatag tgccagtctt gaaggaagtt 120120
acagcctaat aaaagacata gggcatgttg tttggttact gtaatatgaa gtggcatgtg 120180
ttaaatgtca ggggagaact acaaagtcat aaaaaggtgg ggtagtgttg acctaggcct 120240
aggaatcagg aatgacacca tcgggagtaa ggtagtgttg acctaggcct ttaagataca 120300
atagggacag tatggaaaga gtatattttt cccacttaaa ctctttcctt ggtcgttccc 120360
tcaaatttc ccttttgtcc atgtgcaggc actttagtga gtttctgcga agtcaccatt 120420
tctgtaaata ccagattgaa gtgctgacca gtggaactgt ttacctggct gacatctct 120480
tctgtgagtc agccctcttt tatttctctg aggtaaagtc tgcatttctt ttcacatctct 120540
attcgagcat tccagcctct aactatcaat gctggggccc tgtctatagg aaataacaca 120600
gaagagccaa gtcatttcca aaaagatgta tcattgtttc aagttgtttc tgatggcaag 120660
agtaatttaa taatatatta gagagaacat gaaaattcaa tgtattaaat aactctaatt 120720
ttgagaaacc taattaaact actgcatgta agagagtgca tgttttttat tatttggagc 120780
tattttaaaa ccacagaatt tgaaacttgc ttccagtgca taaattgcag accagacttc 120840
agaagagaaa aaaagtagta aattttttct tatgctcatc attttttactt tagtcacttg 120900
ataggattgc ccagtgaaga agcatttgca acagacaatg agtatattaa tcttttttgag 120960
gcatacagtt tagtataatg ctctttgtta ggcttcaaca agtgaaatta ttttgttgga 121020
aagcaaatga ctattaagta gaaaggaggat tcccagctctc acaaagcagt aatttagaca 121080
ctcgattctg cctctttaca agaatacagg tactcagttg atttgtttttc tcactccctt 121140
tctttgctat aagtttaaat caacaatttg tttaggttaa tatgtcctca tggaatggtg 121200
```

44

```
gaaatgatca gatataaaat atttggtttg gttagtttac tctttatatg tttgctggca  121260
aggaaccaca aatccagttt agtataattt ttactctagt tcactaaaag tttgcatcca  121320
gctgtgtagg tagtgtttgt ttcttgttaa cttttttttc gtctaaaaga atactttaaa  121380
acttttcaat ctcaaatgac tgtaacttgc tgacaggtgt taacagaaga agtagatctt  121440
tttgtttttt gcttatgacc tgtattttaa tatttgagct tatagattag agattgtgag  121500
agaaatctgt ttatagtctt attttcccctt gtgtattttc tcttcctagt acgtctaact  121560
agaggatgca gtgaatatct tacaattctg gttggcagca gataacttcc agtctcaget  121620
tgctgccaaa aagggccaat atgatggaca ggaggcacag aatgatgcca tgattttata  121680
tgacaagtga gttatattga tagatggatt cagcagatac ttattgaaca ttgatatgt   121740
tttgtggaaa taaagatgaa taaactcagt ctctgttgtc aaggagctca caggagggcag 121800
cataaaagct gctttttatat ggtgtttgta aagctttggg ggtcttaga acaaaagttt  121860
ctgctgggaa aggggaggtg tatgtgggat aacaggatg gcaatggtgg tgttcaagga  121920
gtgtttccca gaagagagat tttgtttgga tcccaaagaa agaagggaat tttgctaccc  121980
agagaaggca gaaaacaaca ttctaggcaa aggcattggc ccagaagcca tggaaacgta  122040
ggggaaagtg gcactttcaa gaaacttgag ttagataat caaaggagtg gggaataaat  122100
atgaggatgc tggtactaat tggaatagat tgtaaggac cttgaatgcc tatttatggg  122160
tatattatac tttctgtata aatctgctca ggcacgtgt taattagttt tttattagtt  122220
ttcactgaaa atagagggat ggaaacatca tacagtaaac aaaattgaaa atatctggtc  122280
aggcagatga tgagcttgtg gccagctctg taacgtatgg tattctttttc atttaacttt  122340
tcttactctg taaaaaaagt aattcgtggt cgggcacggt ggctcactcc tgtaatcaca  122400
acactttgag aggacgagag aggtgaatcg cttgagccca ggaatttgag accagcctgg  122460
gcaacatggc aaaccccgcc tttactaaaa atacaaaaat tagctgagcg tgatggcgtg  122520
cgcctgtttgt cctagctact taggggcctg aggcagaagg atcacctgag ccttgggagg  122580
tcgaggctgc agtgagctgt gatccactgt actccacccct gggcagggca gtagagtgag  122640
accctgctctc caaaaaaaaa aaaaaacaaca aaggtaattt gttatttgta tccttaagca  122700
aatgctaaag gggtaacttg gggatagaga aaagtccaca gatgttaggg tttgaagaca  122760
ctaatagtat ctaggccagt ggttcctgaa cattagtctg tgggctcttg ctgggctgtc  122820
tgcataggaa tcacctgaga gcttattaaa aatagttttt caggctggtt gcggtggctc  122880
acgcctataa tcccagcact ttgggaggct gaggcaggcg gattacttga ggtcaggcgt  122940
tcaagaccag cctggccaac atggtaaaac cccgtctcta ctaaaaatac aagaattagc  123000
caggcatgat ggcacacacc tgtaatccca gctactcagg aggctgagga aggagaattg  123060
ctcgagcccg ggaggtggag gttgcagtga gcgagatca tgccactgca ctccaggctg  123120
gctgacagag ggagactctg tctcagaaaa aaaaaaaaaa ataggttttc agtctgggta  123180
ccggtggctc acacctgtaa tcccagcact tggggaggcc aaggcaggca gatcacttga  123240
ggtcaggagt ttgagaactg cctggccaac atagtgaaac cttgtctcta ctagaaacta  123300
caaaaaatta actgggcatt ttgacgggtg cctataatcc cagctactag ggaggctgag  123360
gcaggagaat tgcttgaacc cgggagggag aggactgcat ctcaaaaaaa aaaaaaaaa  123420
aaaggtttcc agtcccctg tctcagaaat tctgattctg caggtttgag gtgtgaccag  123480
gaatctttat ttttagaaga cataccagat aattctgata aatagccagt ttagggatgt  123540
agtctaattt tcctattttg caagtaagga aaataaggcc cagagaggta atgatttttct  123600
caaagtcaca gaacaagtta gtggcagaat tttggactgga atgcagttct taagttgtct  123660
tccagtgttt attctggtac agtatgtttg tagaaggtat tacgtaagaa acattgttat  123720
atagatgttg agataggaag agtttacatt tagaaatttg gtctaaaaatg cctgaacatt  123780
caagtcgtgg aggagtattg accaacttac tcaatacaac ataggagatt cacatttgt   123840
tacaaaaatg ctgatttaaa aggagagttt tctttttttt cttctttttt atttttttgag 123900
atggagtctt gctctgtcac ccaggctaga gtgcagtgac acgatctcag ctcactgcaa 123960
cctccacctc ctgggttcaa gcggttctcc tgcctcagcc tcctgagtag ctgggattac 124020
aggtggggc caccacgccc agctaatttt tgtattttta gtagagacag ggtttcacca  124080
tgttggccag gccggtcttg aactcctgac tgcccagcct ccacccacca ctgcctccca  124140
aagtgctgag attataggcg tgagccactg cgcccagcc gcttgttttt gtatcatata  124200
tatgcatcat cataatcatg cattatcaac cttctgtatt ctgtcaggac atagaaaacca 124260
ttagagtgct tggaagagag ccttttcttt tttctgcat ttaatgcttt ttttggtatt  124320
catttcataa tcagcttacc aaaacattac ctgcattata ccccatcaag gtagaaatct  124380
tcccaaatagg tcatatgcat ctagctcacc cctcagtgct gttttgtttt gaatttgtac  124440
atgtttactc ctcgctttt aggggaaaatg accatgtctt ccttcctat aaattcctttt  124500
ctatctatca agtcctcaac agagaatagg taccccataaa tatgtgattg ttagtttctt  124560
tgcctcagtt gtagtctgat ccttacagct tttaaacaac agtagagttc accgtcaaga  124620
actaaggatg gttggcaggc agatagaaag gtagcaagtt gacccaacta tctctgggga  124680
agtgggaaca aagaaaggtt acatcagcac tgtcatcaca tagctctata gttctaggcc  124740
tgcaggctca atcaagtagc cttgtataag attctctgga ggaggtgctg aaagttgctt  124860
atacttgcta tggaattgta ttttacttcg gatatcttt taccataggt acttctccct  124920
ccaagccaca catcctcttg gatttgatga tgttgtacga ttagaaattg aatccaatat  124980
ctgcagggaa ggtgggccac tccccaactg tttcacaact ccattacgtc aggcctggac  125100
aaccatggag aaggtaaccc agaacttcaa acgtatcaaa ctacaagaag ttttattggt  125160
agaactcata aaatataagg tgggaaaacc aagcagaata gcacagtgga aattgaagca  125220
```

45

```
gtccagcaaa gtgattaaga gcagaggcct tgagtctggc ctggtatgta cagtcacgtg  125280
ccacataaca ttttagtcaa cagtggactg cgtgtacgat ggtcctgtac gattataatg  125340
gatcaaagct ggtagtgcaa taataacaaa agttagaaaa aataaatttt aataagtaaa  125400
aaagaaaaaa gaaaaactaa aaagataaaa gaataaccaa gaacaaaaca aaaaaaatta  125460
taatggaagct gaaaaatctc tgttgcctca tatttactgt actatacttt taatcattat  125520
tttagagtgc tccttctact tactaagaaa acagttaact gtaaaacagc ttcagacagg  125580
tccttcagga ggttccaga aggaggcatt gttatcaaag gagatgacgg ctccatgcgt  125640
gttactgccc ctgaagacct tccagtggga caagatgtgg aggtgaaaga aagtgttatt  125700
gatgatcctg accctgtgta ggctaggct aatgtgggtg tttgtcttag tttttaacaa  125760
acaaatttaa aaagaaaaaa aaaattaaaa atagaaaaa gcttataaaa taaggatata  125820
atgaaaatat ttttgtacag ctgtatatgt ttgtgtttta agctgttatg acaacagagt  125880
caaaaagcta aaaaagtaaa aacagttaaa aagttacgat aagctaattt attattaaag  125940
aaaaaaattt taaataaatt tagtgtagcc taagtgtaca gtgtaagtct acagtagtgt  126000
acaatcaagt gctaggcctt cacattcact taccactcac tcgctgactc acccagagca  126060
acttccagtc ttgcaagctc cattcatgt aagtgcccta tacagatgta ccatttttta  126120
tcttttatac tgtatttta ctgtgccttt tctgtatttg tgtttaaata cacaaattct  126180
taccattgca atagtggcct acgatattca ttatgtaac atgtgataca ggtttgtagc  126240
ccaaaagcaa taggttgtac catatagcca aggggtgtag taggccatac catctaggtt  126300
tgtataagta cactctgtga tgttagcaca atggcaagca gcctaacgga aattctgttt  126360
attgattgat tgattgattg attgattgag acagagtttc actccattgt ccaggctgga  126420
gtgcagttgc acagtcttgg cacactgcaa cttctgcctc ccaggttcaa ccaattactc  126480
tgcctccatc tcccaagtag ctgggattac aggcaggcac caccatacct ggctaatttt  126540
tgtattttag tagagacagg gtttcaccat tttggccagg ctgttctcga actcctgacc  126600
ttaagtgatc tgcctgcttt ggcctccgaa agtgctggga ttacaggcat gagctaccat  126660
gcctgggcag taactgaaat tctctaatgc cattttcctt atctgtaaag tgacgataat  126720
atgcacgttt acctcaaagt tactttgatg attaaagtaa ggtaatgtat ataaaataca  126780
tattaacata gtacctgaca catggtaagc atcaaaaaat gttaactact tttattacta  126840
ttattattac gtatttttaa ataattagag agcagtatca aaaattagct gggcgtagtg  126900
gcatgcacct atagttccag ctactcagga ggctgaagct ggaggattgc atgagcctgg  126960
gaattaaaag ctgcagtgag ccgtgttcat gccctgcac tccagccttg gtgacagagc  127020
aagaccctgt cttgaacaat taaagaaggc attatgccgc aacgttagct tagaaatgac  127080
ccacatatat caccagtaac tgtcaacagg attggaaccc tagttttggg tattatgatc  127140
acaaggtatt attaatagct tattaataat aaagcgttgg ctaggcacgg cgactcacat  127200
ctgtaatccc agcactttgg gaggccgagg tgggtggatc acctgaggtc aggagtttga  127260
gaccagcctg accaacatgg agaaaccca tctctactaa aaatacaaaa ttagccgggc  127320
gtggtggtgc atgcctgtaa tcccagctac ttaggaggct gaggcaggaa aatctcttga  127380
acccgggagg cagaggttgc agtgagccga gatcgcacca ttgcactcca gcctgggcaa  127440
caagagcaaa actccgtctc aaaaatataa ttataataaa taaataaaag taaagtattg  127500
atgtttgtga atgatttatt cttctaatga actagaggag attttttccag gaatttcaga  127560
gccagtgagg ttatgttgct tgtatgtgtc atgtgtatcc aggtgaaaaa acttaattaa  127620
acgctattat ataataccat acataaaaac tgaattttag gaatactgaa gaatgacata  127680
tagaagtcaa atcattaaat agctagtagt aaacagaata gagtgtcagc tgttacccaa  127740
tgatgataat attttcacga ttaaaattaa accttttctg attttaaagg aaaagttcag  127800
atctgtatca tataaagaat gtaaattttc agggtaataa aattaaaatg cagagagaaa  127860
aatgcaaaaa tagttcttac tagatgtgtg tatgtaagga acttagacta atttaagaa  127920
cactgtcaag accctggtag ttaggtagga aaaaagacat gaatgattca ttcaacaaaa  127980
actttgagta tttctgtgct agatggtagt gttacatgtg taaacaaaat aaatgtgttt  128040
ctgctatcct ggagcttagt ctacaaaaaa ggtacatatt ggccgggcac ggtggctcac  128100
gcctgtaatc ctagcactt ggaagatcga gcgaggtgga tcacctgagg tcaggagttc  128160
aagaccagct tggccaacat ggcgaaaccc cgtctctact aaaaatacaa aaattaactg  128220
ggtgtggtgg cggcaccctg taatcccagc tactcgggag gctgaggcag gagaatcact  128280
tgaacctggg agacagaggt tccagtgagt cgagtcatt ccactgcatt ccagcccggg  128340
gacaaaagc gaaaatacgt ctcaaaaaca caaaacaaa caacaaaggc acgtattaaa  128400
tacgaacata aatatttaca aattatactg aataagttct catgtttatt atttgctttg  128460
ccagttacaa actttttcct cgtagaatta gaaatataaa taataaacat gaagactcat  128520
tcagtataat taataattat taaatgtaaa taaaaacatc tatgtacaat taggcattta  128580
tttaagaatt atttgaaaaa aaaacaatgt ggaaacagat attttgatat attgctagtg  128640
attgaaatta ataatgttct tttgaaggagt aaagtgacca tatatattaa agttaaaatt  128700
taactcagca atcacacgcc tggtgagtta tcttaaggaa atcagtttga aagtaaaaatc  128760
aatatatgca caaagacttt aacatttatc ataaaccaga aaaatcgagt ttcaaattat  128820
atcctatgca ctatttttctg ctaaaaagta ttaatatcaa cttttatgtaa tactttcgtg  128880
acaaatattt tgggggagaa aacccaacaa aattacatgc attgtaattt tttttttttt  128940
ttttttttta gacagtcttg ctccagccgtc caggctcagga caattctcct gcctcaggcc tcccgagtag  129000
tcactgcaca ctccatctcc caggttcaca caattctcct gcctcaggcc tcccgagtag  129060
ctgggattac aggcgctcac caccatgcct agctaatttt tatagtttt agtagagatg  129120
gggtttcatc atgttggcca ggctggtctt gaactcctgg tctcaagtga tccgtctgcc  129180
tcggcctcct agagtgctga gattacaggt gtaagccact gcacccagcc ttatgcatta  129240
```

46

```
taattttaat ttgtaaactg tacaaaggga taatacttgt agtacaacaa gaagtaaaaa 129300
catttgttat aggtagttaa catttgtaac cagtagaatt ataggtaaaa tttatttatt 129360
taaaacagtt ttagttggat ttgatttcaa ctttaaaata atgcttttca tctctatcag 129420
gtcttttgc ctggcttttt gtccagcaat ctttattata aatatttgaa tgatctcatc 129480
cattcggttc gaggagatga atttctgggc gggaacgtgt cgctgactgc tcctggctct 129540
gttggccctc ctgatgagtc tcacccaggg agttctgaca gctctgcgtc tcaggtattg 129600
actgattgcg tctgccatta gggagaaaag catacacatc ctttccttca catcccagta 129660
acagatccta ttatttgtaa attttaagtt gtggaaaaaa aagataaaag ccaggcacag 129720
tggcctgtgc ctgtaatccc agcactttgg gaggctgcgg tgggcggatc acacgaggtc 129780
aggagttcga gaccagcctg gccgacatgg tgaaaccca tctctactaa aaatacaaaa 129840
attagccggg catggtggca ggcacctgta atcccagcta cttggagagc tgaggcagga 129900
gaatcgcttg aacccaggag gcagaggttg caatgaacca aaatcacgcc actgcactcc 129960
agcctacctt acatatcttct aatcaaagca tttgtggtaa cttaaaatat actgtattgt 130020
aaagtatcat gctgtttcat ttaggccatt attctatttg aatctgtggc tgtttctctt 130080
aataaatcaa gtaatatgga atatattcat agcctctgaa gagctcttta tgtaagtatt 130140
tatttaggat acttttttgta aaataagtga atgaattctt aggtctcctt ttttttttctt 130200
ttcttgagac agggtctcct cgctgcaacc tggaaattct gggctcaaat aatccaccca 130260
ccacagcctc ctgaatagct gggactagag gcatgcacca ccacgcctgg ctaatttgaa 130380
attttttttt ggccaggcat gatggttcac gcctgtaatc ccagcactt gggagaccga 130440
ggcaggcaga tcacgaggtc gggagatgga gaccagcctg gccaacgtgg tgaaaccccg 130500
tctctactaa aaatacaaaa attagctggt tatggtggct catgcctgta atcccagcta 130560
cttgggaggc tgaggcagga gaatggcttc aaccagggag tcggaggttg cagtgagccg 130620
agatcacgcc actgcactcc tgcatggtga cagagttaga ctccatctca aaaaaaattt 130680
ttttttaaa tgatggagtc ttgctgtgtt gctcaggctg tgcttgaacc cctgacctca 130740
aatgccgcct gcttcagcct aagtttcttt tttttttgta aagagacagg gtcttgctat 130800
gttggccagg gtagtctcaa actcctggct tcaagcagtc ctcccacctt ggcctctcaa 130860
agtgctggga ttacaggcgt gaaccactac ctataatgtt gtgtttcact caaggccttt 130920
tgatttcgtt ttgcattacc gtgccacatt gtgcatttcc ttgacctttt ttgggttttt 130980
tggagtgctt tcatatgtta aaccataact gattctcctc aaaatcacac aaagtagaat 131040
atcctaagac aagaaatcta aggaggcata aagaagttaa ctggttttat taaactcaca 131100
cagtaaatga tagagccaga aatattcccc ttctagtgtt cttcaccatc agcttaatgt 131160
agcataataa tttctaatt actgttgaca aataaataac ccttgaatt ttcaatactg 131220
ggccttggat aaattttcct aatttgtaag agagtattat cgtattgcca tttacaaagc 131280
tctcctgagt atcttttttct tctgttaagt ttacctagga gataaactgc tagtagggt 131340
tgccattttg gtttttttgat ataggttaga atgtcttggt tttttttttt tttttttttg 131400
gtttttgttg ttgtcattgt ttgagacagc atcttgctct gtcgcccagg ctgagtctgg 131460
atggcacgat cgtggctcac tgcaacctcc acctcccggg ttcaagcaat tctcctgcct 131520
cagcttccctg agtagctggg attacaggca tgtgcaacca cacctggcta attttttgtat 131580
ttttagtaga gaaggggttt caccatgttg gtcaggctgg tattgaactg ctgacctcat 131640
gatccacctg cctcggcctc ccaaagtgct gggattgcag gcatgagcca ctgcacctgg 131700
ctgaatgtct tgtttttttgat taggcactta agaaaggcct aggtactaac cataaaatat 131760
attttatac cttttgttga tactatatat atagaaaact gcacttatca taacctttaga 131820
caccttgaag aatgttcaca agcagaacta acccatgtga cccagcatcc agatcaaaaa 131880
cagcattatc agccccctcta gaagccctct ccattcactg tccttcttgt 131940
catttataaa taaaaccata ctgtgtattc ttttcttga cagctttatt gtgctaattc 132000
acatttacat catacaattc agtggttttt atatgttcac agagttaggt aacattacc 132120
atcccccact ccaccagccc taggcagcca ctagtctact ttttatctct atagagacaa 132240
tagatttgat tattctggac atttcataaa catggaaccg tatattatgt ggtcttttgt 132300
tgccaactgt ctttcactta gcatcatgtg ttcaaaagag catcatgtta tccatgtttg 132360
atttttatttg aatttttatt cctcattatg gccaaatatc ccattgcaag gatttatgac 132420
gtgtacaagt tttttgtgtg atacaggttt tcttttttgtt tttaaatttg aggtgggagtc 132540
ttgctctgtc gcccaggctg gagtgcagtg gcacaatctc ggctcactgc aacctctgtc 132600
tcctgggttc aagcagttct cctgcctcag cctcccgagt atctgggact ataggcacgc 132660
accaccacgc ccagctaatt ttttagtaga gatggggttt caccatgttg gccagtctgg 132720
tctcgaactc ttgacctcaa gtgatccacc catctcggcc tcccaaagtg ctgggattac 132780
aggggtgagc cactatgccc ggctgtggtt ttcatttctt ttgttgtata tacataggag 132840
tagaattgat gagtcaagag gtaactctta aacttattga aaaactgcca gattgtttttc 132900
cgaaaaggct gcaccatttt gcaatcccac cagcagtgta tgagtttttac agcttctcca 132960
catttcattg gaacttatta tctgtttggc tgtttttaaa aatgatagtc attccaataa 133020
gttctacttc agtgtggttt ttgcacttct ctgatgagta atgatgttga gcatctttttc 133080
atttgcttat tggcctttgt tctagctttg gaaaaatgtt tattcaaatc cttttggccat 133140
ttttatttttt attttttattt atttattttt ttttgagacc aagtctcact ctgtcagcca 133200
ggctggagta caatggtgtg gtctcagctc actgcaacct ccgcctcctg tgttcaagtg 133260
```

47

```
attctcctgc ctcagcctcc cgagtagctg ggattacatt tcaggcacct gccagcatgc   133320
cgggctgatt tttgtatttt tactagtgac agggtttcac catgttagcc aggccgtcca   133380
caaactcctg acctcaggtg atctgcctgc ctaggcttcc caaagtgctg ggattacagg   133440
cgtgagccat tgggcccagc ctagattttc tttttttctt tttttttttga gaaggagtct   133500
tgctcttgtt gcccaggctg gagtgcaatg gcacaatctt ggctcactgc aacctctgcc   133560
tcctgggttc aagcgatttt cctgcctcag cctccccagt agctgggatt acaggtgcct   133620
accaccacac ccagctaact tttgtatttt ttttagagac agggtttcac catgttggcc   133680
aggctggtct caactcctga cctcaggtga tccacctgcc ttggcctccc gaagtgctgg   133740
gattaccggc atgagctacc aggcccagcc aattttctca ttatattgcc caggctggtc   133800
tcaaactcct gggttcaagt gatcctcctg ccttggcctc ccaaagtgtg gggagtacag   133860
gcgtgagcca ccttgctcag ccccttttgcc cattttaaa ttagattgcc ttttatatt   133920
gagtttcagg agtcctttat atattctaga taaatgtccc tatcaaatt atattatttc   133980
caggtatttt cttcattctg tgagttgtct ttcctctacc ttttaaaaaa gatgggtttt   134040
tgtttgtttg tttgtttgtt tttttaagat aaggctccat tctgctgccc aggctggagt   134100
gcagtggcac aatcacagct cactgccact tcaacttcct gggccgaagt gatcctctta   134160
cttcagcctc ctgaatagct aggggccatag ataacacacta tcacacccag cttttttttt   134220
ctgtttgtag agacagatct tactgtgttg cccaagttgg tctcaaactc taggctcaaa   134280
gtgattctcc cacctctgcc tcccagagtg ctgggattac aggtgtgagc cacacgcaac   134340
ggcacccacc accatgcctg gctaattttt ttgcattttt agtagagaca gtgtttcacc   134400
atgttcacc ggctggtctt gaactcctga cctcaggtga ttcacctgcc atggcctccc   134460
aaagtgctgg gattacaggc gtgagccact gcacccggcc aaaatattgc cttcttaaca   134520
gtattgtctt ctaatttgtg aactgatg tatcttcatg tatttatgtg ttctttcatt   134580
tcagcagaat tttgtagttt tcagagtaga agccttttcac ctccttgggt catttattcc   134640
tatgttttaa gttcttttcg attccattat aaatagaatt gttttcttaa tttcattttc   134700
agattgtttg atgagagagc atagaaatac aagtgatttt tacatgttga tcttgcaact   134760
tcaactttga taaatctgat tgttagctct aatagttttc ttgtggattc tttaggattt   134820
tcaatatata agatcatgtc atttatggat agagatagtt tttttttctgg ctagaactta   134880
cagagcaatg atgagtagaa gtggcagaag caaaaatctt tgtcttgttt cctatctgac   134940
agggaaagct ttcagtttca tcatttaata tgatgttagg tgtgggtttt caataaatgc   135000
cttttttcag attcaggaat ttccctatca ttcctgattt tttaaggctt tttttttttt   135060
ttaaatcatg aaaggtgtt gaatattgtc atgttctttc tgtatcagta taaatgatcc   135120
tatggatttt gggttttatt ctgttgatgt gaaatattaa ttgatttca gatgttaaac   135180
caaccttgca tacctgagat gaatctcact tggtcatgt gtataatctt ttcaatatgc   135240
tgctggattc catttactgg tattttgttg aagattttgt atctgaacgc ttaagataac   135300
atttacactc tatcagaaat gaattgacca taaatgtgag agtgtatttg tgggttcttg   135360
attctctctt attccaaaga tagacataca tccgtctgta tgtctgtctt tatgccaagta   135420
ccatctctc ttgattacta ttgctttgta ataagttttg aaatcagaaa gtataaatga   135480
gattttggta tctgagtaac agtcctcata gaattgttg ggaaatattc cctctttatt   135540
ctggtccctc tttctttttt gtttaactgt gtatcttgga gattgttcct tctcaacaca   135600
tgagagccgc tttcctcc ctcccacccc tgctatagag aggtctataa gtgtcgttc   135660
aattattta tttacttaac ctattactta gtcgggaca ttaagcttgt ttatgtcttt   135720
tattttaaac aatgctgcag tgataatct tgtatataag tcattttcca tcaatataag   135780
tctctctgta actgaatttt tagaagtgga atttctaggt caacctatgg ctctgtattt   135840
cacaaaaata ccaattctgg tttttcttgt ggagttgggg agtaggaggt agaatgctgg   135900
aggagaactt gctgtactca gctggctagt cattttagaa aggttcctt agcttctttt   135960
tgtcatatgg cctcacaaag aatcaaaaac attctatct accctgtaaa catggggctt   136020
tactacccaa gatacatatt tctggatgta tgacagcttt tcatattgaa gaaataaatgc   136080
tgtgagtaca gcacatttgt tggaacttag gtcgttaaga atgtcttata aattcataca   136140
ttatacattt tatttttattt tatttttgag actgtaggta atgtcttatca tctgtcgccc   136200
aggccagcgt gcagtggtac aattttggcc aatcttggct cactgcgacc tccatctcct   136260
gattctcatg tctcagcctc cagagtagct atggttacag gcatgcacca ccatgcccgg   136320
ctaatttttt tatttttagt agaaactggg tttcaccata ttgaccatgc tggcctcgaa   136380
ctccttgccc caagtgatcg gcctgcctca gcctcccaaa gtgctgggat ccttgtattg   136440
gagactgagg tgggaggagt cctggagccc aggagggtga ggctgcagtg agttgtgatc   136500
gcgccattgc acttcaacct aggaattata ggcttcagtc actgtgcccg gcatgtacat   136560
tttaatattg tgctttcctc ttttagctat agtatgaagt tacatttcag agtcatttgtt   136620
gttaagcatc ttaatagtga tgaggttgag tgaaagttac ttctatttca aacactgaag   136680
aaaattttgat acaaatctgt cacattccaa gcccaggact gattgtttca tatacttctt   136740
acatatgccg ggtaagctta gctcatgcct agaattttta caagtgtaaa taacttttgca   136800
tcttttaaat tttttaatta aattttacat ttttttctaa tctattatta tatgcccaga   136860
actttcactt agagtgtgca gtataatgtg gtggttaagt ataaaggctc tggagtgact   136920
tcctggtttt taatcttggc tctgccattt attggcagcc gctaacctct tggtatctca   137040
gtttcttcat ctgtaaaatg agaataataa agtgaaaaga tgccaacatc atttactctg   137160
```

48

```
ggctgcataa ctgatacttg gaaaaagtat tcctttgagt ttaagaatta agttggttat  137340
tcattttagc ttgtaataaa aagatagtga ttcataggat atgccactta ctgaaattta  137400
ccacagatcc aatcataaaa tcactttctc ttccctaaag atagcttgat taacatgtaa  137460
aggtgtgtaa aggcttgatt acactaccct gatccgtacc ccagttccca gcagcaccat  137520
gaaaaaggga tttcaacata tttaattact ttcagtagaa agtaacagtg gtaggccag  137580
cgcagtggct cacacctgta atcccagcac tttgggaggc cgaggtgggc ggatcacgag  137640
gtcaggagat tgagaccatc ctggctaaca cgatgaaacc ccgtctctac taaaaataca  137700
aaaaattagc cgggcatggt ggcaggcacc tgtagtccca gctacttggg aggctgagac  137760
aggagaatgg cgtgagcccg ggaggcggag cttgcagtga gcttagattg tgccactgca  137820
ctccagcctg cgcagtggag cgagactctt gtctcaaaaa aaaagaaagt aacagtgcta  137880
ttgggagact gaggagccta gaaagtactt gaaggaagta aaagtttgt ttgaccacat  137940
tgtatttgga aagccagctt tttcagctgt gtcagcttg tgtagtgatt tttagttctt  138000
ctttttagaaa ataaccggaca aggccgggca cggtggctca cgacctgtaa ctgcacctct  138060
tgggaggccg agacgggcgga attacctgat ctcagagtt cgagaccagc ctgggcaaca  138120
gtagtcccag ctactcccgga ggctgaggca ggagaattgc ttgaacccgg gaggcggagg  138180
ttgcagtgag ccaagatcac accattgcac tgcagcctgc gcgacagagt aagactctgt  138240
ctcaaaaaat aataataaaa taaaaaaagaa tggacagtaa aacctaaatga gttcattccc  138300
aaagatgatg tttattcttaa gggatggttc atttatttaa gaccttacat aaagtctatc  138360
aattgcgtga tttttcacttt ctgtaattgt gtgtatgtat aatgtaaata tatatgtttt  138420
tgttttgttt tggttttttg agacggagtc tcgctctgtt gctcaggctg gaatggtggc  138480
gtgcaatctc agctctctgc aacctctgtc tcccaggttc aagcgtttct tctgcctcat  138540
cctcccaagt agctgggact accaggacgt gccaccacgc ccggctaatt ttttgtattt  138600
ttagtagaga tggggttttca ccgtgttagc caggatggtc tcaatctcct gacctcgtga  138660
tccaccccgcc ttggcttccc aaagtgttgc tattacaggc atgagccacc acacccagca  138720
tgtattttttt aaatgtataa aatgaagcag aaaagagaaa tgataatttt tcttcatctt  138780
gaaagattat cttcaccagg cgcagtggct cacacttgta atcccagcac tttgggaggc  138840
ctcggcaggc ggctcacttg agttcgaaac cagcctggcc gacatgtga aactccgtct  138900
ctactaaaaa taaataaata aagatggttt taatatatgt tttagtttta tgattttagc  138960
atcttttctga aattttttctc aaggcaagta aatttgtatc agttggtata ttggtaccca  139020
tctatgaaat aacttattag gaagatatct ctaaaataag atcactttgc ctaaaataaa  139080
ctgatatatt gatgttcaca gaattttttct tttaaccgac ttgataaatg cattattctt  139140
gacgtcaagt gatccacctt cctcagcctc ccaaagtgct gggattacac acatgagcca  139200
ccgcacctgg cattattctt ataaaaggtt aaattttctag ttaagtttaa tgtcctcttt  139260
gttcatgtac cattgcttat tttcttcct tcctactcac agtaatcatt cttatgtgtat  139320
gcacttttgt ttgcttattt ttatgtaatt gatattacgc tccattctgt acgttgtact  139380
ttcattcaca gtgagttttg gacattccta tgttcattca tacagactta cttcatttta  139440
actacactgt agtattccgt atgtaatatt tactataact catcactgta gcagagcatc  139500
tcatagtgta tgtattactg ttttgatcatt ttggtatcaa tgagtattta agtcattttg  139560
agttttttccc tcttataccc agtattacag aggatctctt tttatatgct tctttgtacc  139620
aagaggcaga ttaaaaaatt tttttttgaa aaaattttttg aaaaaaaatg aaatgaagtc  139680
tcactatgtt gcccaggctg gtctcaaact cctaggctca agcaatcctt ccatcttggc  139740
ctcccaaagt gctggggatta caggcatgag ccaccatgcc tggcctacat tttaaatttt  139800
gatagctctt acaatttact ttgtaaagta tctgcatcat tttatgttct caccagtctt  139860
taataagaat acttcatact tttggctgga cacagtggct cacgcctgta atcccagcac  139920
tttgggaggc cgaggcgggc agatcaagag atcggacggac cctggccaa tatggtgaaa  139980
ccctgtctct actaaaaata caaaaattag ccgggcgtgg tggcgcaccc gtagtcccag  140040
ctactcgaga ggctggagca ggagaatcac ttgaacccgg gaggtggagg ttgcagtgaa  140100
cttagatcac accactgcac tccagcctag cagtcttaag tgtttgctaa tgagattgag  140160
aaaagaatac ttcagactta attctttctc cagtcttaag tgtttgctaa tgagattgag  140220
tttctttttgg tatgttctct gatgttcag gtttttttctt ttatgaattg actgttcatc  140340
tcttttttcac attatttctg ttgggtgatt ttattagtga cttgttaaaa ttctgtatat  140400
tttttcagca tgacactttca ttattcaaaa aaaaaaaaag attctctatg tttctcgata  140460
ctaatcattg gtgtgtaata ccttaaaaat aagacccttta ctgtatttttt tgcttttttt  140520
tttttttttt tttttttttt tttgagatag agtcttgctc tgttgcccag gctggagtgc  140580
aatgqtgtga tctcggctca cagctcactg caactgcaac ctctacctcc ctgtttcaag  140640
caattctcct gccttagcct cccaagtagc tgggattaca ggcatccacc accaccacca  140700
gctaattttt gtatttttag tagagacagg gtttcaccat gttggccagg ctggtctcaa  140760
actactggcc tcaagtgatc cgcctgcctc ggcatccaca agtactggga ttacaggcat  140820
gagccacagt gcctagccac tttttgcttt ttaacttttgt tttatagtac tatagttttta  140880
gtataaacagt atgtatgtat acacacaact atggcttttat aatatgtttc agtcatttgtt  140940
agagcaaggc ctacctttttg ggtgcttctt ttacaaaatt gtcttggctca ttcttgtgcc  141000
tttttttctta tttgtgaatt ttagaattgt gaattacctg ttgactcacc atgttttttgta  141060
aactgagggat tttgaatgga attgcactca attacaagatt atcttgcttt ctgtgcagca  141120
atgtttttatt tcaaataatc cctacttttaa attacttttagg atagctataa atttgtgttttc  141180
tggctttttcta gatttttagatg aaacgcttta aattgattgt ttttctcctaa atttaaaact  141240
gattgttaga agttaaagtc ttctgttcat tcttattttag gaagatgaca tttggaagag  141300
```

49

```
tcagtgactt ggggcaattc atccgagaat ctgagcctga acctgatgta aggaaatcaa    141360
aaggtttgtg gtgtttttat acttcatatt aagcctttac tcacattagt gattgactgt    141420
aagtcaaaga ccacttaagg tttaaactgt ttattttgta aagtaaccac tgtatctttc    141480
accttgtgtt tatagtcaga agtaagtaca agggcttcct gtagtcacat ctttatgcaa    141540
tctcctctga atcaaaagtt agtgaacttg ctttgccact ccgaaaggca catgaatatg    141600
aaaaagcatt gtctattttc ttatttaatg gcaaaataCC cgacctaagt tggacttaat    141660
gtttgagacc gtttatttta ttaaattata ttttttctct tttctttttt tttttgaga    141720
cagttctttgc tctgtcaccc agaccggagt gcagtggtct gaccgcacct cactgcaacc    141780
tctgcttcct aggttcaagc gattttcctg cctcatcctc ctgagtagct gaggactaca    141840
gtgcgcacca ccacacctgg ctaattttg tattttagc agagatgagg tttcaccacg    141900
ttggctaggc tggtctcata ctcctgacct caagcaatcc atccgccttg gcctcccaaa    141960
gtgctgggat tacaagtgtg agccaccatg cctggcctta ttaattatt ttattaaat    142020
ttcctcaaga ttgatgaaag taatgaaata taaagtaat gaaatatatg tgaacataat    142080
tgagttcatg tcctttgtag ggacatggat gaagctggaa accatcattc tgagcaaact    142140
gtctcaagga tagaaaacca aacaccgcat gctctcactc ataggtggga attgaacaat    142200
gagaacactt ggacacaggg tggggacact cacagctggg ggcctgtcgt ggggtggggc    142260
gctggggggag gaatagcatt aggagatata cctaatataa atgacgagtt aatgggtgca    142320
gcacaccaac atggtacatg tatacatatg taacaaagct gcacgttgtg cacatgtacc    142380
ctagaactta aagtataata aattaaaaa aaataaatat atgtggaaaa tattaatagg    142440
tcaaaattca aattgttcat ttaatcagaa gagtagttta gtcaaatcca agggttagac    142500
aacagaaatc tttttgtca agtgcattct ttgtgactga tttcattttc ttcctggttt    142560
acacagagag atttcagaaa caaatgtgga tccgtgacag atggtatcta gaagttttta    142620
gtttggttga attgacagta ttttattgag taaaagatac taattttgt aagaagaaaa    142680
attcaatttt gataagtatg tttaagatta agagctattg gccaggcgct gtggctcatg    142740
cctgtaatcc tagcactttg ggaagctgga gcaggtgggt cacgaggtca agagattgag    142800
accatcctgg ccaacatggt gaaaccctgt ctctactaaa ttagccaggc gtggtggcac    142860
atgcctgtgc accgcctcc gggtttaagc gatcctactg cctcaggctc ctgagtagct    142920
gggattacag gcgccatggc taattttgc atttttagta gagacagggt ttcactacat    142980
tggccaggct ggtctggtct caaactcctg acctcaggtg atctgcccgc cttagcctcc    143040
caaagtgctg ggattacagg catgattcac catgtctggc catttatctt atttctttt    143100
ttttttttt ttttgtttga gacggagtct tgctgtgtcg cccagactg gagtgcaatg    143160
gtgcgatctc agctcactgc aacctctgcc tcctgggttc aagcaattct cctgcctcag    143220
tcttccaagt agctgggatt acaggcgcgt gccaccacat ctagctaatt tttgtatttt    143280
tagtagagac aggtttcac catgttggcc aggctggtct cggaactcct gacctcgtaa    143340
tctgcccacc tcggcctccc aaagtgctga gattacaagt gtgagccact gtgccagcc    143400
atcttatttt ctttcttttt ttttgtcggg tgggagggggg acagagtcta gctctgtcgc    143460
caggctggc tcactgcaac ctctgcccc caggttctag caattattct gcctcagcct    143520
cccaagtagc tgggattaca ggcgccgcca accacgcctg gctaattttt tgtatttttt    143580
agtagagatg gggtttttgct atgttgacca tgctggcctc aagtgatccg cccaccttgg    143640
cctcccaaag tactgggctt acaggcgtga gcttgtattg ggtaaaagaa caatattgag    143700
ggctgcatgg tggttcatac ctgtaatctg agcactttgt gagactgaga tggaaggagt    143760
gttggagccc aggaggggtga ggctgcggct gcagtgaatt gtgatcacgc cattggcactt    143820
ccacctaggt aatgagagcaa gaccatgtct ctaaaaaaca aaacacaatt tttttaagga    143880
atactgggaa gaggtcagtg gtggttttag aacagaggaa gtgccagatg accttggtga    143940
ggcattggcc aggaagaact ctacagtgtc ttaggtagc ttctgtccat aaggataatg    144000
gggtctcctc cccagtatta atagaaaatc ttaggaaggt ttttttttgt ttgtttgttt    144060
tgtttttttt tcctgagatg gagtctctct ctgtcggcca ggctggagtg ctgtggacgg    144120
atcttggctc actgcaagct ctgcctccca gttcacaca attctcctgc ctcagcctcc    144180
caagtagctg ggactacagg tgtcccaccac cacgcccagc taatttttg ttattttta    144240
tagagatgag gtttcaccat gtcagccagg atggtctga tctcctgacc tcgtgatccg    144300
ctcgcctctg ccttgcaaaa tgctgggaat acaggcgtga gccacgtgca ctggcctggt    144360
ttttttgttg ttgttattta tttattttatt tatttttttt ttgagacaga tctctgctct    144420
gtcgcccggg ctgagagtag ctggtggatc atggctcacac gccaagctct gcctgccagg    144480
ttcaagcgat tctcctgcct cagcctcctg agtagcaggg accacgtgcc ctccgccacca    144540
cgcccggcta atttttgta tttttagag acaggtttca ccatggctgc ctggccaggt    144600
gtctcgatct cctgatgtcg tgatccgcccc acctcggcct cccaaagtgc tgggattaca    144660
ggtgtgagcc accgtgcctg gccttgatttt tttttttttt taatctggtc tcatacctct    144720
gacgacttcat tatgtatatg tgttagcata gtgttaacat    144780
agcataggtg ttcggtgttt gcagtttctg tttgttttat atgaattaag gtgtattatg    144840
agcagtttgaa gatatatgga aatttttttc ccaaaccact atctctgctc gttctattca    144900
ttcagtctgt ttatgttatt ccttcattca ttcattttat agaacagtgg agtgcctact    144960
gtatgcatct attgttctgu gtcctgggga agaaaacaaa gttcctgctt tcatggaact    145020
tacattatat tggcggagac agtaacagac aaacaaaatgt agctgttga catgtgttac    145080
atgaaaagca gggtaggggg ctgggagaga gtagtaggga gtgctatttt cgaggtggtt    145140
gtcaggaaag gcctcactga ggaggtggca ttttgagtag acctgagcgc agcggggggcg    145200
taagcccagg cagcatgtgg aggaagagtg ttcttggtga aaggaacaag gatagaggcc    145260
                                                                     145320
```

50

```
cgaagctaga gagctcagca tgatcaagga acagcaagcc ccgtgtggct ggaatggagt   145380
gagcaaagga atgagcagta gaaggtgagt gagttgggag gtcaccagag accatggcaa   145440
ggacttgaaa gtgtcaggga cacattggaa gttggagcag ggaaatgatg ggatttatgt   145500
tttgtttttg ttttatgttt agtgttttta agggattgct ctatcagcta tttggaaaat   145560
ttagtgtagg gcttcaagaa gagaagcaga gaaacaacat tcttgccata gtcatatgct   145620
aagtaaggga tgatggtggt gtggattagg ctggtagtgg aagaccagtc cagttcgggt   145680
tgtatttgaa ggtagaggca aaaagattat atttctacca gcaagcccat ctatgaagtt   145740
acttgtatta ttaatttaat tgagacatgc ccacataaac taataaaatag gaatttctgc   145800
agtttggtta aacaccoctg tatatcctgg ttcttctttt agttgtctca gatgtctct    145860
aagtcaagta tttttttggtg gtgtaggagc ctagagattg aatttattca cccaaaaggc   145920
atttgagtga ttactatgtg ccaggcacta tgctgaatgc caaggatgta aataagaggg   145980
cgtagtctca gtctgtttta ctccagcttg gttcctttt tcctggttgt gttatttagc    146040
catatcagtt atcctacaga atgtttaatc ttctgtactt tcctggttgt gttatttagc    146100
ttatttctct ttccttgaca tttcttgtaa actggaagtt acacctatag tcttgatgat   146160
tcgtgttaca cattttagat tagaacacat catgtgttgt atatggtgtt tttgaaagcc    146220
tctctgtata ttggtctgta cattaaaatg ttgcctgaat ggatacacat aaaatttaac   146280
agtgattaca ttagagatga gaagaaagag gtgccttttta cttttcaata tacctttcc   146340
tctgcttttt gaactttctt gccctatgca tacgttattg cttaatcatc cacctcatct   146400
cttcccctgt ggctttctgt tgcatttgga atgaaatcta gcctctttgc tgttacctgt   146460
ggatgtcccct tgctggcctc tatcacctta ctttgaacca ctcctttcat ggactgagct   146520
ctcattggac tatctttat tcttttgctg aagtttcttc acttttgagtg cctctgcagt   146580
tgctatttca tggctgtggc aagccctgcc atggctttca tgcaaggatg gttcctcctt   146640
ctcatctcaa tattatctct tcagagaggg accttcccaa ctccgatgat ctaaaatcct   146700
ttgtatatac cactcactac cacttcttttc tttcttttc ctttttatctt ttttttttt    146760
tttttttttt gagataggggt cttgctctgt tgcccaggct ggaatcacga ctcactgcag   146820
cctcatcttc ttgggctcaa atgatcctct cacctcagcc tctcgagtag ctggaactgc   146880
aggcacacac caccatactt ggcttattat tttactttt gtagagacag ggtttcacca    146940
aggctggtct caagctcctg ccgcaagcaa tccacatctc tcagcctccc aaagtattgg   147000
gattatagga gtgagccact actcctggcc tattttctta ttcactgtct aaaattatct    147060
tgttcattta tttacatact tgtttatagc ttattttctca gctggacatg gtgcctcaca   147120
cctgtaatct caatactttg ggaggctggg ttggagaatt ggttggaccc aggacttcaa   147180
gaccagcctg ggcaacaaag tgagaccctg tctataaaaa attgtttaaa aattagctgg   147240
gcatggtggt gcatgcctgt ggtcccagct acttgggagg cagaggtggg agaatcgctt   147300
ggccccagga ggttgaggcg acggtgagcc atgattgtgc cactgcactc tagcctagtg   147360
acagagtgag accatgtgtc taaaaagtaa ataaaaaatag ttttctcttc atgactagaa    147420
tattacctct atgtgggcag ggagtttgtc tatactattt ggcactatat ttcctgattc    147480
tgaaattatg cctagcacat ggtaagtact cctttaaatat ttattgactg aatatttaa    147540
tacttaagaa tttcatttgg gattatctga gtggtaagat tacggattat atttatgtaa   147600
gaaaaaatca tttttttaac ttggttgccc tttgccacac tgacatagac actaagtttt   147660
cttagccaga ttacttccga ggatactcac agaggccatt ctcttctcaa tccccaaata   147720
attgatattt cttagcactt tcaagctaat gcaattctta gatgatgtat ctgtgtatat   147780
catatcctca ttctacaaat gtagaaattg aagtctgggc acagtggctc tcacctgtaa   147840
tctcagcagt ttgggaggcc aaggcgaggc gatcactgag gacaagagtt aagacccagcc   147900
tggccaacat ggtaaagcct tgcctctatt aaaaataaca caattagggc cgggcgtggt   147960
ggctcacgcc tataatccca gcacgttggg aggccaaggc aggcagatca cgaggtcagg   148020
agttcgagac catcctggtc aacacagtga aaccccatct ctactaaaaa tacaaaaaat   148080
tagccaggca tggtggcacg cgcttgtaac ctgtctcaaa gggaggctga ggcaggtgaa   148140
tcccttgaac ccgggaggcg gaggttgcaa tgagctgaga ttgcaccgct gaactccagc   148200
ctggtcaaca gagggagact ctgtctcaaa aaaaaaaaaa aaaacaatt agccaggcgt   148260
ggtggcgggt acgatacct gtaatcccag ctactaggga ggctgaggga ggagaatcac   148320
ttaaacccag gaggtggagt ttgcagcgga ctgatgaatgc accactacat tccagcctgg   148380
gcaacagagt gagactctgt cttaaaaaaa aaaaaagaa agaaagaaat tgaggaatgt   148440
ggagattgtg gtctgtgatt tgtaaggaat tgtttacag ttcggctccc cttcctcctgc   148500
tttggttcag aatacccct tgacaatggt ttgtttacag ttcggctccc cttcctctgc    148560
ctttctctcc ttcctatctg agggcagctg gaaagaattt tcatccattta ctagcctata   148620
gctttaattt gagttttgaa accttgataa tagagcacag aggaaaagac tagttttct    148680
tttttgaga cagtcttgct ctatggccca ggctggagtg cagtgacacc atctcagctg   148740
gttgcaacct ctgcctcccca ggttcaagca atttgcctc agcctcccga gtagctgaga    148800
ttacaggcac gtgtcaccac gcccagctaa ttttctgttt ttgtttcgtt ttgtttttt    148860
ctgagatgga gtcttgctct gtcacccagg ctggagtgca gtggtgcgat gttggctcac   148920
tcaaacctct gtctcctggg tttcaagcaat tctttgacctc agcctcccga gtagtgggg    148980
actacaggta cgtgccacca tccctagttc attttgtat gttttagtaga gacggggtttt    149040
cactagtgtg accaggctgg tctcgaactc ctgatctcag gtgatctact cgtctcagtt   149100
tcccaaagtg ctgggattat tggcacacgc ctattttttgt attttagta gagacggggtt    149160
ttaccatgt tggttagact ggtctcaaac ttctgacctc aagtgatttg cccgcccaag   149220
cctcccaaag tgctgggatt acaggcgtga gccaccgtgc ccagccaaga ttgagttttg   149280
aaaagagcct tctgagatta tgagaagggc aagcaagata acttaagaag ttacattaaa   149340
```

51

```
atcatctaag agacagtgta acaagaagga attgtaaaat gatgttatga gcacgtgccc   149400
aatgtagtgg caatcccttg tgcttcgata cattggtggg agacaaaact gtacttaaat   149460
tgataaatcc cttacatgtc attttaagga gcttagactg actcccatca tgtagacatc   149520
agagatttct tttttttttt tttttttttt tttgtgacag agttttgctc             149580
ttgttgccga ggctggagtg caatggcgtg atctcggctc accacaacct ccacctccca   149640
ggttcaagca attctcctgc ctcagcctcc cgagtagctg ggattacagc catgcaccac   149700
cacgcctggc taattttgta tttttagtag gacgggggtt tctccatgtt gttggctggtc   149760
tcgaactcct gacctcaggt gatcctcccg cctcagccac ccaaagttct gaaattacag   149820
gcgtgagcca ccgcgcccag cccagagatt tctaaacaga gttctaacca gatgcttttc   149880
cctgtcagta gaatgagaat gaattgggag ggagaagaaa aagattcaaa gttaggtagt   149940
agccagtgga attagctggt aatgttgata ggagagagaa aagattcaaa gttaggtagt   150000
ggtagcaaga attagaggga aggtcggatt tatgatatgt ccaaggttga attctaaggt   150060
gaaatttggt ggcagatttc atgtgtaaat tgggaagtga gattgagttt ttttaacatg   150120
ggtttctaa catgtcaata gagtgactct gcaggggggc ctgacgagag aacagtgcat   150180
actctggttg gcagtaaaat ttcattaaac caatatttaa acccttaggt aataataaaa   150240
attgagggaa aaggatccag tgtctttgtatt ttttatgaat tcagttattg aattaaacag   150300
gaccttgcct caagaaataa tctaccaaca attaacttgt tttaaagcaa agttaggaag   150360
tgagcatgtt caaattatta aataaaaaag taagtggtgt atttcattca tagaaaataga   150420
ggctggccta cttcgggatga ttctcagcat gtgattacag atgtgggctt atacatccta   150480
gggagttaag gcgtactctg gcttggatag agtagagctc tttgaaactc ttctctcacc   150540
cagctagttt atatagacta gagaactaga tagtagcagc atactctgtc ttagaagccc   150600
tttatatag gagctggtct gtgaaggtttg aaaacataac aaatgtgttg gtgtctcca   150660
atgtattgct agattcttac ccaagagcat tatcctggtt agggtttggt ttggtttttgt   150720
tttgttttt aatgtttgcc acaaactaac actagatgtt agttctttca tcaagtgagg   150780
agagtagaag aaaagtccag aactctgaaa caccttttca aaagtttttc aagccatgat   150840
gtttgcaagt taaatgctct gttatgtaag caatataatc agtttttatt aatgtaacat   150900
tccttagtgt tttggggtat cacacaaaaa agaatatcca tatctggaag caacagcttt   150960
taaataagag cattgtggtg gtggtggtga tagtggtttt ttttttttt tttgagttgg   151020
agtctcgctc tgttgcccag gttggagtgc agtggcacga tctcagctcg cttcaacctc   151080
tgctcccagg ttcaagcaat tcttctgcct cagcctcctg agtagctggg attatggca   151140
cctgctacca tgcctggctg attttattta ttttagtaga gacaggtttc accatgttgg   151200
ccaggctggt cttgaactct taacctcagg tgaatcaccc acctcggcct cccaaagtgc   151260
tggaattaca ggcatgaacc accatggcca gccaaataag agcattttta atgtaaaatt   151320
atgcatgaaa tgtacattca attttgtctt tgtttactag gatccatgtt ctcacaagct   151380
atgaagaaat gggtgcaagg aaatactgat gaggtaaatc ctaccttag gataaaaaga   151440
tttctgttta taagtgccac cctcatgtaa gtgaggttta aattttcct tttctttagg   151500
tcccatgttt aagcagcatg gcacatttat gttctcttac ccagaatgta ccaagaaagg   151560
gtggtccctt cttaacatct aacaattgcc tggtagtagc agtgaaggta tcttcagtca   151620
gaggctagga ccactgaagg atatacatgc attcaagttt ccatcagcca gcaggcatca   151680
gtaatcagtg tgtagatcaa aagctcaaat gtttccttcc ccactggcag ttttacttca   151740
agtagtggag gcttgctttt ttaatagtta attaagtaca ttgagagatg ggagtgaaa   151800
aaaggaaaat gttttatttt gaccatctaa tatgaaagta gttcgggtgt aggtatccag   151860
tagttgacac tggaagacag ggaatgacat gttaatattc atagccagag ggtgggcccag   151920
gtttttttgt acatgggaat gaaattctta tccaaataag tagaaattat gtgcgtaagc   151980
catttgttaa gagcactgag tatgtgcatc tcgatccatc taatgaataa ccattatcac   152040
cagtttaaat tattttcttt aggcccagga agagctagct tggaagattg ctaaaatgat   152100
agtcagtgac attatgcagc aggtcagta tgatcaaccg ttagagaaat ctacaaaggt   152160
aaggatgact tcgttttgtg taaatcaaaa agtattattt tccaggtgta aaaataaaaa   152220
agaacataag gggttttctt gccttttgaag gattaactgc tgtgggatt accttcttat   152340
cataaggac tagaaaattg acaaactaaa tgaaacaact gtttgcatat attggacaat   152400
gggcaataca gggaaccaat ggaaacacaa cacaggcccag tagtcttgct gacaagaaa   152460
gttaaatatc aaagttcagg ccaggtgcag tggctcacgc ctgtaatccc agcactttgg   152520
gaggccaagg cgggtgaatc acttgaggtc aggagttcaa gaccagcctg gccaacatgg   152580
tgaaaccctg tcttagccgg gtgtggtggc aggcacctgt aatcccaact atttgggagg   152640
ctgaggcagg agaatcgctt gaaccaggga ggcggaggtt gcagtgagcc gagatcacac   152700
cactgcactc cagcctgggc gacgagcgaa accccattc aaaaaaaaaa tcaaagttca   152760
gagagctcaa tttgagtgaa agttgtagga taaggtagca gaaaagagga agctgcccag   152820
aaagaaagcc gtagagatat ttagagagat tcccatggat ccttggccta ggagtgatct   152880
gtatatgtgt ggggtgaaaa cgcatgtgtc caggtagaga accccccaga aattagtagg   152940
ctgaatgatt gctggaacat agggctaaga aaagttcatg gccagaagga tctggccaga   153000
gtagagagac ttagtaatac acaaggcatt gggtagtgtc ttcacagagg ttatgcctta   153060
ctactgaaga taaattagtc ctagagtaca agcacctgaa ccaagtttca aagcaaattt   153120
ttaaaggtgtc aaattaccta acaactgcat gccaaaacaa aggcctaacc ctctttacag   153180
taacacaaca aaattcagca cttcacagtg taaagttaga atgtctgacg tccaggctgg   153240
gcgcagtggc tcatgcctgt aatcccagca ctttgggagg ccgaggcagg tagatgacct   153300
gaggtcagga gttcaagacc agcctggcta acatggtgca accccgtctc tattaaaaat   153360
```

52

```
acaaaaactt agccaggcat ggtggccggc acctgtgatc ccggctactt gggaggctga    153420
ggcaggagaa ttgcctgaac ccaggaggtg aaggttgcag tgagccgaga tcgcaccact    153480
gcactctggt ctgggcaaaa agagcaaaac tcaggctcaa aaaaaaaaaa gaatgtctga    153540
cgtcaatcac aaattaccaa gcatgacatg aagttgacct ataaccagga gaaaactcaa    153600
tctatagaaa cagacccaga tgtgagaaag atgatgaatt tagcagacaa agaccatcaa    153660
gtggctattt taaatattaa aaatatgttc aagtggccag gtgcagtggc tcatgcctgt    153720
aatcccagca ctttgggagg ccaaggtggg taggagttca agaccagctt ggccaatatg    153780
gtgaaacccc ttctctacta aaaatacaaa aaaattagct gggcatggtg gcaggtgcct    153840
atagtcccag ctatatggga ggctgaggca caagatcac ttgaaccggg gaggtggagg    153900
ttgaggttgc agtaagccga gattgtgcca cttgtactcc agcctgggca acagagtgag    153960
actctgtctc aaaaaaaaaa aaaaaaaagt taaagaaaac aagagtataa tgagaaaaat    154020
gcaaaatagt tttaaaagaa ccaaatggaa tttcttaaaa taaaaaatac cagaaatggg    154080
ggccgggcgt ggtagctcac gtctataatc ccagcaactt gtgggggctg aggcaggcga    154140
atcacctgag atcggtagtt caaggccagc cctgatccca tggagaaacc tcatctctac    154200
taaaaataca aaattagctg ggcgtggtgg cgcattgcct gtaatcccag ctacttggga    154260
ggctgaggca ggagaattgc ttgaacccgg gaggcagagg ttgcggtgag ctgagattgc    154320
aaaacagtag actcgaagaa ctagctgagt ttttctttac tttaggcagt aagtgtgacc    154380
ttttgcaggt gactactttа gttcctcatg tcctcattag tagatcagag aaattcgaca    154440
ccaaaacccc aaaagaaaaa cccttctaa tcctcattcc atgattttat gaatgcatga    154500
agtcctaggc ctgcgaagga atactcattc tctttatcct gtgtttgataс ctctctgctt    154560
caacctccaa ctcgacattt gcctatagga tgtacttgga cattcagcat aaactacctc    154620
acaccattac tgaattgctt catgtgcaca tgtcccatgc cacaataccg gggaaccttg    154680
cttccgtgat atttgtccgc agtgctgtga ctacaggagg gagtcagtga atgtctgcat    154740
gtgtgtcttt accatccctc ttgaatatgc tctagggtta attcctagaa gtagaattac    154800
tctattgaaa attggcaata tttttcattc taatatctat tgccaacatg ggaaagcaag    154860
tctggatgcc agtccttgtt atatgcccct tgggtaagtt acgtaacctc tttaagcttc    154920
tgttcactca tattttaaca aggaaaatta caatattta cctcacaaaa ttgtagtcag    154980
cttctggctg tcttaaactc tggtatatag taaacactaa gtgtttggtgt ccatccttaa    155040
tttgtaataa taggtcactt gttagagaaa tgcaccttac catttttctt tcttttcttt    155100
tttcagttat gactcaaaac ttgagataaa ggaaatctgc ttgtgaaaaa taagagaact    155160
tttttccctt ggttggattc ttcaacacag ccaatgaaaa cagcactata tttctgatct    155220
gtcactgttg ttttccaggag agaatgggag acaatcctag acttccacca taatgcagtt    155280
acctggtaggc ataattgatg cacatgatgt tcacacagtg agagtcttaa agatacaaaa    155340
tggtattgtt tacattacta gaaaattatt agttttccaa tggcaataac ccatttatga    155400
gagtgtttta gcctactgga atagacaggg accacatcct ctgggaagca gataagcata    155460
gaactgatac ttgatgcaca ctcgtagtgg taactcatcc ctaatcagca ttgtaaagca    155520
ggtgccagag gtggtttgct ttgtccttcc aaagcaggtg agtcagcccc accgagagcc    155580
aggcagcttt gagtggcagc gtggtgctag cagcttcagc ggaacagggt gagagttaat    155640
tatgcagtct tcttgacagc ggcattaatt tggaaggaaa ctgacaagtc atgggtcaag    155700
tttcagtgac ttcctccttc ctctgatggc agtatatagt tttcacattt taattcctcc    155760
tcctgagatg cactatactt aaaaccattc tctcccctgc taacagaagg gtgtgaatct    155820
ggtttacttt gagcattagg atttgcccct ttggaattct gcactccagc ttttgccctt    155880
tcccttcaga atacatgtgg aaagaaagaa agaaatagcg agaaataggc tacttaactt    155940
gtggcacctt gaacaaagca gttcttccca aattatactt ttttttttt taaataaggt    156000
gagcaggaga actgggggaga gagaaacatt ttctttctcct tctttttagga tagtaagaga    156060
ctcactcact gcacttctgc tcagttggct tctgcatggg gatcacacag ccatcagcag    156120
gactgccctga ttggtgagca cactccattg accacgtgac gccagcgctt cctcaatgca    156180
catgattgag aggaaagaaa gtctctctag atgttactgc ttttgctcag actttgcaaa    156240
aaaaaaaata tatatatata tgtataaata tataattatt aatcactttt gtccttgaga    156300
aagtcttgaa tgaacagaga atttattcca ttgcaatatt tgattgtata gaggcacact    156360
gttttcatcc tttaaagctg aagtattcat gtacttaaac catattatat ttaaatatgt    156420
ctgttttaa aatatatata tatgaaatcc atttaaacca atcattttt gtcctttgaga    156480
gggcatttaa ggctcttctg ttgaaataaa tatgaaatct atttaaaatt gtgtcaactt    156540
ctaaaaaccc agaacatgca ccactactgg actttgcctt gtgtttgaag tgtatggcat    156600
aaacccaagg tttttattag tcatctatgc tgtgattaat tcattttgtt ctttaacaa    156660
aatatttcca tccacttcac attgcttcaa tcttaacag aaaagcaata taaaggttat    156720
agaataaaat gtggttttgg gcaactcttg ctgcctctgc atgttttgga ataacaattt    156780
ctacaagact ctaggctgtt taaactagtg ctttcagtta agataaaatttc taatcatttc    156840
tttgtatata cattttgtgc ttctgagcta gagatgccaa gtagttgtaa actgcttata    156900
aagagaatag cagcaaattt gagactcggc tacttttttc tgccccacct gcttgagac    156960
acagaagcgg agtgtggccc gaaattatta gccagattta atatttgatc taaagtaggt    157020
ccttgtactc attttaaagt tggaatttga ttcctccaac attgagcacc caccatgttc    157080
caggctctgt gcattgtgcc cacaaaataa gattccctgg tggagttttt atggggttcaa    157320
ataatcagtt gaacaccctt catctttatc atgttgttga cattgacaca aattgtttaa    157380
```

53

```
aaagaaaaga tattagagag aaagtggtac ctttgtaact tgatgtgtct tcatcattcg  157440
gtaagatttg atgaaagtaa aaagcaaatg tcagccaaat ccagtgaaca gcaataaaac  157500
agggagtaac tttttataac tttttctact tggatttcaa cattcagtag agcttttcga  157560
aatgtaagta gtttacagta ctggaggttt gactagttca gtaggaattt ggaggggaag  157620
gtcattctga attgtaacaa agtacaaact tctttgctgt tttatttaag tactgagagc  157680
taagcacctg atgaagtgac tgacctctct ccagtgacag tgtttgggta cctgcctgac  157740
ttcaggagtg gggtttatgt ttctacacag tgaccttttc tctcgccctc tcctccctct  157800
tgcccacaca ccagttgatt ggacctgggt tgaactcctg atccagacag gcccaagaca  157860
gttcttaatg ttaagaattt tggggccggg cacggtggct catgcctgta attgcaacac  157920
tttgggaggc cgagacaggc ggatcacttg aggtcaggag ttcgaggcca gcctggccaa  157980
catggtgaaa ccctgtcttt actaaaaata caaaaattag ctgggcatgg tggcgcacgc  158040
ctgtaatccc agctacgtgg gtggctgaga caggggaatc gcttgaacct gggaggcgga  158100
gttgtgcaat gagccgagac cgtgtcactg cattccagcc tgggtgacag agggagactc  158160
tgtctccaaa aataaaaata agaaaaagaa ttttgggcta ggtgcagtgg ctcacgcctg  158220
taattacagc attttggaag gcccaagatg ggcagatcac ttgaggacag gagttcgaga  158280
ccagcctgga caacatggtg aaactccatc tctactaaaa agacaaaagt tagccagatg  158340
tggtgatggg cacctatatt cctagctcct cgggaggctg gggcaggaga atcacttgaa  158400
cccaggaagc agagattgca gtgagccaag atacatctc tgcactccag cctgggcaac  158460
agagcaagac tctgtctcaa aaaaaaaaga attttggccag gcgcagtggt tcacgcctgt  158520
aatcccagca ctttgggagg ccaaggcagg cagatcacga ggtcaggaga tcgagattga  158580
cctggctaac atggtgaaac cctgtctcta ctaaaatac aaaaacattag ccgggtgtgg  158640
tggtgggcac ctgtagtccc agctactagg gaggctgagg cagaggaagg atgtgaaccc  158700
aggaggcgga gcttgcagta agccaagatc gtgccactgc actacagtct gggcgacaga  158760
gtgagactcc gtctcaaaaa aaaaaagaat tttggccggg tgcggtggca catgcctgta  158820
gtcccagcac tttgggaggc caaagtgggc ggattacctg aggtcaggag ttcaagacca  158880
gtcccggaac tatggcgaaa ccctgtctct tactaaaaaaa aatacaaaaa ttagccaggt  158940
gtggtggcgg gcacctgggg aggctgaggc agggagaaat gcttgaaccg gggaggcgga  159000
ggttgcagta agccaagatc gtgccactgc actccagcctg aagactcttt ctcaaaaaaa  159060
aaaaaaaaag aattttgcat ggggaaggag agatactgtt caccatctgg aatggtgctt  159120
ggatgtggca cttacaaaat caggagccag cactgcatgg acaaacagaa gcatgtgggc  159180
ctgagatagc aggtaccttg ataacctga agacatcctt ggtttctgca tctattcctg  159240
catccttgca ttggactaca ttaatctgtc agttatcctt ataatgattt ttgattttttt  159300
ttttttgaga tggagtttcg ctcttgttgc ccaggctgga gtgcaatggc acgatctcgg  159360
ctcaccacaa cctccacctc ccaggttcaa gtgattctgc tgcctcagcc tcctgagtaa  159420
ctgggattac aggcatgcgc caccacacct ggctaatttt gtatttttag tagagacgga  159480
gtttctccat gttggtcagg ctggtctcga actcccaacc tcaggtgatc acctgtctc  159540
ggcctcccaa agtgctggga ttacaggcgt aagccatggt acccggtctg tttttctggt  159600
ttttgaaacc agtctgaagt gagtttttttt aattacgtga aaggagtttg gctaaaatac  159660
tgccatactg ccctaatgcc taatgattat gtattctcag catgtctgca aatggtgctt  159720
gatttctgga gaataatttt tctttagtaa acttcactta agtcgtcatg tgtattctct  159780
caaaatggta tcctaaccta atggagctaa aagacacccc ttgttttttat aacaagcagt  159840
tactgaggcc caggaagggg agaagtccct ggcttgtgag atgatcacca ttagaactca  159900
ggcctgggcc agtgcctttt catgcttctc agatccttcc aaagaataat gaagattata  159960
accgcttttta gcaattgtaa taaacccaga aatagaaagc tttttggtta gagtactggt  160020
agaagttttgg cgggagagat aattttttaca aaattgtaa ataactgcca attctatata  160080
ctaggcaagg tctctggcct tgtaaaaccc ttcaaggtta caactttggt ggcccacact  160140
aatagttacc cactgaggcc ctctccgggt gaacattgag cactagagga agcccctctg  160200
cttgggcagg actgggcgtg gtgcagagta ggagcggtga tactgtgat tctgggcagg  160260
tggagatgac cagtgatgtc caataaagga cactggaggg agcagtgtga gtaaaggccc  160320
tgagggcatt catgttcagg gaggggttgct gcccactggc ttgcttgcca cacaggagag  160380
tgggtattcc tgcctagta actttatgta aacaagtatt tcctcagtct gttcctctca  160440
aactgcctgc tctggcacat tcagaatgtc acagaactca cctggatgca ttccagcccct  160500
tgcctaaagg tgacagtgca tctccttccc caccccaccc ctcataccac tcaagcacct  160560
gtcagactgg cccagtctgt gggcaaggag cctagagagg gcttagtttc agcttgaaag  160620
gagctgggat ttaccaagaa gcaaatgaga gacgaggatt gcaacaactg tgccatttcc  160680
ccagcttcag ctgactcctg tatattgact gtgccttcag actcatccgt aagtgaccccc  160740
aggctggcct ctcccacatc acagtaagaa ttccacacac catacaactt ggaaagaggc  160800
tccagctgaa ggaagcccca cacttctttc aagtttttct tagtcttctc ttcttggcaa  160860
agagtacctt ttgtttcttc taattatgta actattggtt tagtaaatat tcacccattc  160920
agtcaccctg taagtggcag gcactgttta cagggacaca ggaaggaata aaaacttgca  160980
ggcaccttgg agcttgcatt ctattgaaga ggtaatggaa gttgggatag cagctaaact  161040
atgctggtat tggccaggcg cagtggctca cacctgtaat cccagcactt tgggaggcca  161100
ggtgggcaga tcatgaagtc aggagatcga gaccatcctg gctaacatgg tgaaaccccg  161160
tctctactaa aagtaaaaaa aaaaattagc caggtgtggt ggcgggcgcc tgtagtcca  161220
gctacttggg aggctgaggc aggagaatgg tgtgaacccca ggaggcgaag attgcagtga  161280
gccgagatgg caccactgca ctccagcctg ggtgacagga cgagactctg tctcagaaaa  161340
aaaaaatatg ctggtagttt tgattcaaga tggcccatga gcccatgat ttaggtctcg  161400
```

```
tacccaccaa ggtctactgg aaaacatcag gctctcctgc tatagaccca tagggagagc   161460
tgcagccgag aggggagct gaagagaagt gcccccttctg tgtcctgtca gcctcatcct   161520
tccgcaagga ccagttgctg tgccactcca ttcacttgct gcaagactgg aggtttttcc   161580
tcaggtgttg agcacctggt ttacaagatg tcagcatctt gatgcctgag accatcaagg   161640
caagtctctg aacagggctt accttagagt aaggcttaga agaggccgta aagtcagtct   161700
cagctccgtg gctctgcaga gctttgggac atgtgaattc ttaaaaacaa gactattgta   161760
cagttactat atgcatgcag tataaaatta taaccttgga aaatcctagc tagctgttga   161820
gctaattcca taaagtaatc agctcctgag ttctgcagtg gtaataataa tcagcataat   161880
gagtaaaacac tgtgtgtgcc aggcagcgtc tcatttgatc cttgtgataa tcttgtaagt   161940
actgattttc tccctttctt aaacaaagtt ttttttttt tttttagag ggtctcacta   162000
tgttgcccag gctagtcttg aattc                                        162025

<210> 36
<211> 162025
<212> DNA
<213> Homo Sapien

<220>
<221> mutation
<222> 156,277
<223> Nucleotide Base Change: T to C

<400> 36
gaattcctat ttcaaaagaa acaaatgggc caagtatggt ggctcatacc tgtaatccca      60
gcactttggg aggccgaggt gagtgggtca cttgaggtca ggagttccag gccagtctgg     120
ccaacatggt gaaacactgt ctctactaaa aatacaaaaa ttagccgggc gtggtggcgg     180
gcacctgtaa tcccagctac tcaggaggct gaggcaggag aattgcttga acctgggaga     240
tggaggttgc agtgagccga gatcgcgcca ctgctctcca gcctgggtgg cagagtgaga     300
ctctgtctca aaaagaaaca aagaaataaa tgaaacaatt ttgttcacat atatttcaca     360
aatttgaaat gttaaaggta ttatggtcac tgatatcctg tttcattctt tatataatca     420
ttaagtttga aatgtatact tgcactacta acacagtagt taatcttagt cctacaagtt     480
actgctttta cacaatatat tttcgtaata tgtatgcact ggtgtttatg tacgtgttta     540
tgtttatatc tgttaaaatt agcagttttcc atcttttct atttttgtacc atacaactaa     600
ttcagaagga ttgacagagc aaaatgattt gatgaagtat aaaagtcaca tggtgagtgg     660
cataaataca actctgaaca attaggaggc tcactattga ctggaactaa actgcaagcc     720
agaaagacac atatcctata tgtcaagaga tgtaccaccc aggcagttaa agaagggaag     780
tacacataga aagcacaatg gtgaataatt aaaaaattgg aatttatcag acctgggaga     840
catttgctcc taaagtcaga gtcctctatt gtttttttgt ttttgtgggt ttctttttaa     900
attttttat tttttgtaga gtcggagtct cactgtgtta cccgggctgg tctagaactc     960
ctggcctcaa acaaacctcc tgcctcagct tcccaaagca ttgggattac agacatgagc    1020
cactgagccc agcccagacg ctttagcatt tatgaagctt ctgaaatagt tgtagaaacc    1080
gcataagctt tccatgtcac tttcaaagtt tgatggtctc tttagtaaac caaccaagtt    1140
attcctcaag ggcaaaataa catttctcag tgcaaaactg atgcacttca ttaccaaaag    1200
gaaaagacca caactataga ggcgtcattg aaagctgcac tcttcagagg ccaaaaaaaa    1260
aggtacaaac acatactaat ggaacattct ttacaagagc cccaaagtta atgataaaca    1320
ttttcatcaa agagaaaaga gaacaaggtg ttagcaaatt cctctatcaa atacactaa    1380
acatcaagga acatcaatgg catgccatgt ggaagaggaa gtgctagctc atgtacaaac    1440
cagtagataa tttcaacttg ctgccgaatg aaacctcttt gcaaggtatg aatcagcact    1500
tctcatgttt gttttgcttt gttttgtttt gtttttagag acaggcccctt gctctgtcac    1560
acaggctgga gtgcagtggc acgatcagag ctcactgcaa cctgaaactc ctgggctcaa    1620
gggatcctcc tgccttagcc tcccaagtag ctgggactac aggcccacca tgcccagcta    1680
atttttttaaa ttttctatag agatgggatc tcactagcac ctttcatgtt tgatgttcat    1740
atacaacgac caaggtacaa tgtggaaaag ggtctcaggg atctaaagtg aaggaggacc    1800
agaaagaaaa ggggttgcta catagagtag aagaagttgc acttcatgcc agtctacaat    1860
actgctgttt tcctcagagc agagttgatg atctaaatca ggggtcccca accccccagtt    1920
catagcctgt taggaaccgg gccacacagc aggaggtgag caataggcaa gcgagcatta    1980
ccacctgggc ttcacctccc gtcagatcag tgatgtcatt agattctcat aggaccatga    2040
accctattgt gaactgagca tgcaagggat gtaggttttc cgctctttat gagactctaa    2100
tgccggaaga tctgtcactg tcttccatca ccctgagatg ggaacatcta gttgcaggaa    2160
aacaacctca gggctcccat tgattctata ttacagtgag ttgtatcatt atttcattct    2220
atattacaat gtaataataa tagaaataaa ggcacaatag gccaggcgtg gtggctcaca    2280
cctgtaatcc cagcacttcg ggaggccaag gcaggcggat cacgaggtca ggagattgag    2340
accatcctgg ctaaaacggt gaaaccccgt ctactaaaaa ttcaaaaaaa aattagccgg    2400
gtgtggtggt gggcacctgt agtcccagct actcgagagg ctgaggcagg agaatggtgt    2460
gaacctggga ggcagagctt gaggtaagcc gagatcacgc cactgcactc cagcctgggc    2520
gacagagcga tactctgtct caaaaaaaa aaaaaaaaaa aaagaaataa agtgaacaat    2580
aaatgtaatg tggctgaatc attccaaaac aatccccca ccccagttca cggaaaaatt    2640
```

55

```
ctcccacaaa accagtccct ggtgccaaaa aggttgggga ccgctaatct aaataatcta 2700
atcttcattc aatgctaaaa aatgaataaa cttttttta aatacacggt ctcactttgt 2760
tgcccaggct ggagtacggt ggcatgatca cagctcactg tagcctcaat cacccaggcc 2820
ccagcgatcc tcccacctaa acttcctgag tagctgggac tacaggcacg caccaccatg 2880
cccagctaat ttttaaattt tttatagaga tgggggtctc accatgttgc ccagactggt 2940
ctcaaaccct gggctcaagt gatcctccct caaactcctg gactcaagtg atcctccttc 3000
cttggcctcc caaagtgctg ggattacaag catgagccac tgtacccagc tggataaaca 3060
ttttaagtcg cactacagtc atggacaatc aggcttttca acatgcagta tggacagtga 3120
gtcccagggt ctgctttttcc atactgaaat acatgtgata ctaaggagaa aggtgctcgc 3180
aaggatattt aaaatgaaga atatttaaaa gagatttttt cactcaagta gtatatttca 3240
ataaggctga taaagaccat ttctgtgatc tcaggtgatt cactcaagta gtatatttca 3300
gtaatcatta tctggaacag cctgaatctc aactcaaaata ccatgatttt ttaatgctgt 3360
tatgatacct tgatgtatatg accaaactgc aatgtaggca gctaaatctc cacgagtttg 3420
acttccccga gagttgacag tttttcttcac aaattaaaga aatatatttt ttgatacatg 3480
attggcatat ttaaaaacta cactgaaatg ctgcaaaatg atataaagaa acattttcca 3540
gaatcaaatg caatcaaaga gtggattagg aatctactca ccattatcaa ctaaatagaa 3600
acacttggac tgggtgtggt ggctcacatc tgtaatccca gcactttggg aggccaaggc 3660
aggtggatcg cttgagccca ggagctcaag accagcctga gcaacatagc aaaactctgt 3720
ctctacaaaa aaaaaaaaaa attaaccagg catggtggca gatgcttgta atcccagcta 3780
ctctggaagc tgaagtagga ggactgcttg agcccaggag atcaagactg cagtgagccg 3840
tggtcatgct gcgccacagc ctgagtgaca gagagagacc ctgtctcaaa aacaaaaaca 3900
aacaaaaaac acttaaactt cctgtttttt gctgttgttg ttgttgtttg tttgttttga 3960
gatgagttct caactctgtc cccaggctgg cgtcagtctg ccggtcactg cctcactgca 4020
agctctgcct cccgggttca cgccattcc ctgcctcagc ctcccgagta gctgggacta 4080
tagggcccg ccaccacgcc cggctacttt tttgcatttt tagtagagat gggggtttcac 4140
cgtgttagcc aggatggtct tgatctcctg acctcgtgat ccacctgcct cggcctccca 4200
aagtgctggg attacaggca tgagccaccg caccggccca acctttctgt tttttagttt 4260
gatatgcttg ttaactcagc agctgaaaga atgctgaaag tggccttcag taaaaaaatt 4320
tcactagaat ctctacatcc atatttaatc tgaatgcata tccagattga tcagttagag 4380
caaaaacact catcatcatt cctgatgacc tctaattctg gtttcggctt tctatttcaa 4440
tggaaacaga ataaggaaag aaatggaagg gctctggaaa tttgtcctgg gctatagata 4500
ctatcaaaga tcaccaacaa taagatctct cctataaata taaaacaagt ataattaatt 4560
ttttaattat tttttttctct tcaggagatt ttatttcaag ataaaacata acttctaccc 4620
atactattga ttccaaaggt tagaaaaagt gttttttcctc atcttatcct tcaaagaggt 4680
cacagcaatg caaacatcta taaatgcct ctgcataatt gtcagaagct atagtccaga 4740
aatcattgaa aatgctttc cattttaagc ttagtgagg tgtcttagga aacctctatg 4800
acaacttact ctatttattg ggaggtaaac tcccagactc tcccagggtc tcctgtattg 4860
atctcatttt ttaggcttcc taatcccttg aagcacaatc gaaaaagccc tggatctctt 4920
ttctgcacat atcatcgcgg aattcattcg gcttccagca agctgacact ccatgataca 4980
agcggcctcg cccttctccg gacgccagtc cttgctgcgg ttagctagga tgaggggttt 5040
gctgggcttc agtgcaggct tctgcgggtt cccaagccgc accaggtggc ctcacaggct 5100
ggatgtcacc attgcacact gagctcctgg caggctgtac caattttta attatttaat 5160
atttattttt aaaattatgg tgaatatttt ggtattctgc tctaaaataag gcccataaat 5220
gcacagcaga tatctcttgg aacccacagc tttccactgg aagaactaag tattttttctt 5280
ttaaagatgc tactaagtct ctgaaaagtc cagatcctct acctcttcc atcccaaact 5340
aagacttgga atttatgaga gatctagcta acagaaatcc cagacacatc attggttctt 5400
cccagagtgc agtcctccta aagaggctca gccctaagca ggccctgca ccaggagggt 5460
gggtctgaga cccacatagc acttcccaag gtgcatgtgt cagacagggca ctgaaacagc 5520
tgagcacaga cctgcaagcc tggagaactc tcacagtcag aacggagggg gcccagtggg 5580
actaacataa agagaaaagg gaacacagag aaatggatgg caccaacaac cagcaaagcc 5640
ttcatggcca atgaaagcat cagtgacggg gccagaactc tcatcccaa agactcttca 5700
ctgccttttag tgaaaacaa tggctagaga gtgaagttat gatcatgtat agagaggtaa 5760
agttacattt ttatattctg acttctgtaa tgtgaaattc cctatctgct agactaaaag 5820
tttcagacac cctgttcaaa tatcccatta gttgctagag acttaaaatg aacagaacgc 5880
acattgtcag gatgactatt accaaaaaat caaaagacag caagtattgg tgaggatgta 5940
gagaaactgg aacttttgtg cactgtttat gagaatgtaa aatggagcag ctgctgtgga 6000
aaagagtatg caggttcctc aaagagtaaa accaagatgt ggaaacaact aaatgcccat 6060
cagtggagta aggggtagac aatatgtggt atatacatac catggagtac tattcagcct 6120
ctaaaaaaaa aaaaggaaat tctataacat gcaacagcat ggatgaatct tgaggacatt 6180
ttgctaatga aataaggcag tcatagaaag acaaatactg cacgactcca cttatatgag 6240
ataccaaaaa tagacaaatt catagatcca aagagtacaa tggaggttac ctggagctgc 6300
aggcgggaaa acgaggagtt actaatcaac gaacataacg ttgcagttaa gtaagatgaa 6360
taagctctca agatcagctg tacaacactg tacctagagt caacaataat gtattgtaca 6420
cttaaaaatt tgttaaggat agattaacaa atgtagtaga tccacaaatg tggttaagtg 6480
ttcttaccac agtaaaataa aaaaagaata tcaagcccag gagttcgaga ctagcctggg 6540
taacatggtg aaaccctgtc tctacagaaa atacaaaaat tagccagctg tggaggtgca 6600
ctcctaggga ggctgaggtg ggaggcttgc ttgagcccaa gaggtcaagg ctgcagtgag 6660
```

56

```
ccatgattgc accactgtac tccagcccag atgacagagc aagacaccac cccccccaaa  6720
aaaagaaaaa gaatatcaaa cattttaaaa gatcagatac gcaagaacaa caacaaaaaa  6780
gagatgaaca gagcatcgac cctcatctag tgggattctt ggtctaactg aaaaacagac  6840
attgagagac aaacaatgac agtgatgtga tcacagcaat tacacaggta tcccctgggg  6900
actgcagaag aaaggaggaa tgcctaactt tcagaaaata gagaaagcgt caaacagttg  6960
gtgaaagcct tccaaaacta gagagaactg cacacaccaa atcacagaaa gaacaggaga  7020
cgtgggagat tctgggaccc accggctatt tttgatggct gaacaccctg ctgcaggaga  7080
gacaggagct ggaaagcatg gtgggatgaa acctcaaaca gcttcgcctg cattgcttaa  7140
gatgactggg cttgattaac tctagtcaat ggggacaatt caatcaaaga agaaagatgc  7200
tcaaattcac attttagaat gattttttat ggcagtatgg ggaatagatt aaaaagagagt  7260
gaagctggag gcaagaaact tgttaagagg caactgaaac agtctagatg ataaataata  7320
aactgacaga gtgactagaa aaatcagaac aggctgaatc aacagatacc tagatgaaaa  7380
taacaggact tgatcaccag ttgtatcttg gagagaagg agttgtttcc ttgctttccc  7440
tacgactggg aatacggaag gtttgccgtg tgtattggtt atatactggt gtgtagccaa  7500
tcactgacaa ccatttagca gcttaaaaca caaaggctta tctcccagtt tctgtgggcc  7560
aggaatctaa gataggctta gctggctggt tctggctcag agtttctcaa gaggttgcaa  7620
tcaagatgtc agctggggtt gcatcatctg aaggctcaac tggggccgga gggtccactt  7680
ccaaggagtt cactcacctg cctgacaagg cagtgctggt tgttggcagg agatctcaat  7740
tcattgccaa gtgagcctct ctatagcatt gctggaacat cctccccatc tggcagttgg  7800
ctttctctcag catgagtgat ctgagagaga gacgcaaggag gaagccacag tgttcttcct  7860
actcctactc ctaacactat ggacctactc ctaacactct cacttctgcc ttattccatt  7920
agttagaaag ggaactaagc tccacctctt gaaataagaa gtgtcaaaga atttgtggat  7980
atatttaaaa atcatcacac tgtggaagtg gatagggggt tcaattaatg ctgaagttga  8040
aatgcctgag acattcaaat gtccaacagg caatgaacat acccatagat ggtcatgact  8100
ttagcaagaa tagaggaaga tcacagaatt aaggaggaat tgaaaggtaa aaagaatgga  8160
gtcagattcc ccctgcaaaag tgagccatga aaggaacttt aactattgag ttagaggtca  8220
gagtagcaaa tttcggtgga attcttttttt aaagaaagga accatataag catgttttga  8280
ggtagagggga gaataaatca gtagacaggg agagtaaaa aacatatatg atagggggata  8340
gttgacaaag gtcttggcag aatcccttac ccattgactt ggggccaaga gagggacact  8400
tcttttgtttg agggataagg aaaataagaa agaatgggtg ctatttagtg tggtcctgtc  8460
tctagggcaa acgcataggt aacaaactgt gtgtgttagg aatatagatg tgacctcaca  8520
ttgagattct cacctcaaat ccattttgtt gttacctgta ccttcctacc ttctcttttt  8580
gctacatgca gactgctgtt ttgtcttcct ggcctgttcc aggtttcagc attctggcat  8640
atctgctacc ctgttcccaa acctctctag agtccatgct ccttccttgg atagtgtttg  8700
attgggccac gtatctaaga agtgatgcct tcagttaggc ctgagaacct cctattggcc  8760
aatctccatc agtgaccctg acagacttgg tatcttggag atgtcactgc tcccagcctg  8820
tggtctaaga gaatctcaac ctgggcctct agtagtatgg ataaggcgtt aaggtatctt  8880
tgaaccagag tctgtcatat tcctcaaatgt gggacagata aaacagtggt agtgctgggta  8940
tttctgagct agaactctgg tttttggtct agattctttg atgtatgacc tttcaggaga  9000
attaaaattt gttctaatac aatgttcaat acaaatgtag ttccttttct gttaggacct  9060
caacaaaaca tgaccaactg tagatgaaca ttaaactatg acaattcatg gcacaatagg  9120
cagtaataacc tgcggttccc ccatttttagc agtcactatg gtgacatttg gcacaaatgg  9180
ctatttaagg gtgcttttgt taaaacctac catcttacta ggcacatgat attgaaacta  9240
atgaaataat ggagaaactt cttaaaaact ttaatgaat aaagtgatga agtgataaaa  9300
ttttagctgc tatttataaa gtgactatta caggtcaaac attcttctag ggttttttttg  9360
ttgaagttgt cacatttaat ccttaataac ccactatgag tcaggtattc ttctctccc  9420
cctgcatgta gaaaatctga gatttgtaca ggaacgtatc aaactctgaa gtccatgctt  9480
ctattttccc atgctgcctt tctaataaaa ggtaactaat gctactggat gctgcccca  9540
aagtagagtca ctttcacccc accctacttg attttctcca taaaactaat cacatcctga  9600
caacttattt attgctgatc tcccccacta gattataaac tcaataaag caagatcctt  9660
gtctgctgaa tatcagtacc taaaacgctg tctagcacag agcaagtaat taatatttgt  9720
tgaatgaaca aataaagaaa aaaaattcaa aggaagaaaa agccctaaaa cagatgtta  9780
cctaaacata catttaaaa gaaagcatat aacaaattca ggacagaatt taaatttgat  9840
tttttaaaga aataaccaag tgctagctgg cacagtggc tcacacctgt aatcctagca  9900
ctctgggagg ccgaggcagg cagatcactt gaggtcaaga gttcaagacc agcctggcca  9960
acatggtgaa acctgtctct actaaaaata cagaaattat ccaggcatgg tggcaggtcc  10020
ctgtaacccc agctactcag gaggctgagt caggagaatt gcttgaacct gggaggcaga  10080
ggttgcagtg ggccaagatt gcaccactgc actccagcct gagtaacaaa gcaagactct  10140
gtctgaagga gaggaagaa aagaaggaaa gaaggaaaga aggaaggaaga aggaaggaaa  10200
gaaggaaga gaaagaaaga aggaaagaaa gaaggaagga gaaagaggaa gaaagaaaga  10260
aagaaggaa gaaagaaaga gaaagaagag accaagtgct tatttgggac ctactatgct  10320
atgtttttcc atgcacgcta ttttcagtaa agcagttagc aaacttgcaa gatcataaca  10380
acaaatatat gcttctataa ctctaaaatt gtgctttaag aagttcctct ttaccagctc  10440
atgtatgcat tagttttcta agagttacta gtaacttttt ccctggagaa tatccacagc  10500
cagtttattt aaccaaagga ggatgcttac taacatgaag ttatcaaatg tgagcctaag  10560
ttgggccagt tcatgttaat atactccaga acaaaaacca tcctactgtc ctctgacaat  10620
                                                                   10680
```

57

```
tttacctgaa aattcatttt ccacattacc aaggagccag ggtaggagaa tatagaaaga   10740
ccacccaaga atccttactt ctttcagcaa aatcaattca aagtaggtaa ctaaacacat   10800
gccctaacaa tgaatagcag attgtgctca gaagaatgat ctacaacatc ttactgtgaa   10860
ggaactactg aaatattcca ataagacttc tctccaaaat gattttattg aatttgcatt   10920
ttaaaaaata ttttaagcct aaattttaaa aggtttgata ttggtacatg aatagacaaa   10980
cagacatgga ctagaccaag aattaggttc aaacatatac aggaatttaa tatacgataa   11040
atctagtatt ccaaaggaac caacaaatgg tgttcagaca gcaggatagg agcaccaaat   11100
aacacagttg ggcaccctac cttactccta acaccaggag taactgaagg agcaccaaat   11160
atttatttat tttaattata gttttaagtt ctagggtacg tgtgcacaac atgcaggttt   11220
attacatagg tatacatgtg ccatgttggt gggagtgtcc aaatatttaa aagaaaaaaa   11280
ttggccaggg gcggtggctc acacctgtaa tcccagcact ttgggaggcc gaggtgggag   11340
gatcacgtga ggtcgggagt tcgagaccag cctgagcaac atggagaaac cccatctcta   11400
ctaaaaatac aaaattagcc aggcatggtg gcacatgcct gtaatcccag ctacttggga   11460
ggctgaggca ggagaatcgc tttaatctgg aggacacagg ttgcggtgag ctgagatatt   11520
gcactccagc ctgggcaaca agagcaaaac ttcaactcaa aaaaattaat aaataaataa   11580
aaataaagaa agaaaagaaa aaaatgaaaa tagtataatt agcagaagaa aacaccgtag   11640
aatcctcgga ctcttaggat gggaatgcc tataataaa aaaccctgaa gttataaaaa   11700
agaaaatcac ctacatacaa accaaatctt tctacatgcc taaaacatag cacaaacaca   11760
gctaaataat catagctgaa tgaactggga aaacaaaact tgactcatat ccagacagag   11820
ttaattttcc tacacataaa gagtacctat ataaacccaa caaaaaaacc accactaacc   11880
caaaataaaa atgtgacagg taatgaacag gtagttcaca gagaatacaa atggctcttc   11940
ggcacataag atgctcagac tgactttac ttatttattt tttgagagac agggtctcac   12000
gatgttgccc aggttaggct caaactcctg ggctcaaatg atagtaccag gactacaggt   12060
gtgccccacc gcacctggct cctcaaccac ctgtattaac aggaaatgca aaataaaact   12120
ttcaaatcta ttttacctat tagaatggca aaaatttgaa aaacttcaaa catcatcagtg   12180
ttggtgagaa tgtgaggaga ctggcactct cattttttgc tgatagcata tatatactga   12240
tggcttctat ggaaagcaat ctggcagcgt ctatcaaatg tacaagtgca tatatccttt   12300
gacaaagcaa ttccactcta ggaatgtgtt ctatatggtt gtgcttcctg gggctgggaa   12360
ctgggagcta agggacaggg gcagaagata atcttcttt ccctccttcc ccgttaaaca   12420
tgttgaattt tatatactgt aatatattat ttttcaaaa agataatttt taagcgatat   12480
gtctgggaat tttttttttt cttttctgag acaggtgtct actctgtcat ccaggctgga   12540
atgccatggt atgatctcag ctgactgcag cctcgacctc ctgggttcaa gcaatcctcc   12600
cacctcagcc tcctgagtag ctgggactac aggcacgtgc catcatgctca attttgtat   12660
atacagggtc tcactatgtt gcccaggcta atgtcaaact cctaggctca agcaatccac   12720
ccacctcagg ctccaaagtg ctgggattac aggcgtgagc caccgcgcct ggccctggga   12780
attcttacaa aagaaaaaat atctactctc ccctttctatt aaagtcaaaa cagagaagga   12840
aattcaacct ataatgaaag tagagaaggg cctcaacctt cct gagcaacaaa cacaaaggct   12900
atttctgaga caggaatttg ctgaacaaaa tcgagggaag atgacaagaa tcaagactca   12960
cttctccggct gggcgcagtg gctcacacct gtaatcccag cacttttggga ggccgaggcg   13020
gacagatcac gaggtcagga gattgagacc actactggcta acacagtgaa acccagtctc   13080
tactaaaaat acaaaaaaatt agccgggcgt ggtggcaggt gcctgtagtc ccagctactt   13140
gggaagctga ggcaggagaa tggcgtgaac ccaggaaggc gagcttgcag tgacgcagga   13200
tcacgccact gcactccagc ctgggtgaca gagcaagact ctgtctcaaa aaaaaaaaaa   13260
aagactcatt tctctagatc ttgagccgta ttcaaattta tctcagctta gtgagaggtt   13320
aaagcaagga atatccttcc ctgtgggccc tgctccttac tgaaggaagg taacggatga   13380
gtcaaggaca ccaatggaga aaagcactaa caccattcct tgatgaacat tacgtgaaga   13440
aaatggaata taggaaatcc ataaaagtga ttaaaaagat gttagagcct tcatttgggg   13500
gaccacaggg tcaggagatc gagaccatcc tgcctcaac ggtgaaaccc catctctact   13620
gactgaggga ggagaatggc atgaaccggt gcaggcacct gtagtcccaa ctactcggga   13680
gccactgcac tccagcctgg gtgacagagt gagactccat ctcaaaaaaa aaggttagat   13740
acgagagata aagatccatc agacacacaa cctcaaaaaa aaagttagat gaaagatcac   13800
gcacaggaaa agaaaattca agatataaca ccggaaaact ttcctgaaat tgagtaactg   13860
aatctatagc ttgaaagggt ttagcatatg ccaagaaaaa tcagtagagt ccaaccagca   13920
caagacacat ctagcaaggc tggtgattct accaacacag agaaagaagt gggtgaccca   13980
taatgcggaa aaaggcagac catctgcagt cttctccaga acactggagt ctgaagcaa   14040
aagaatgctg cctactgagc cagaagggag agaaagtgac ccaacaatc tttaccaagt   14100
tagaatgtca cgcattattt aaaggctgca aaagccatga aagacatgaa agaacacaag   14220
catttacaac atgaaagaac acaagcatttc tcatactcaa gaatccttaa gaaaaattgta   14280
gtcctaatcc agcccactga aagttaaatg tacttaatgt gctcattaat gggaacttca   14340
tagcttcaaa tcagtctggt cccatctacc aacatctctc gcccggcttt cctgcaatag   14400
tcagcacctt tccctcctcc cagtcttgtc ccctggagtc tgctctcagc atagcagagt   14460
gaccacatca acacccaagt cagagccctc cagtgcgcac tggtctacaa agcccttccc   14520
accccccacc ccacgtgccc tccggatcct tgtgacgtgt ctcctgcata ccctagcagc   14580
cctggcctcc tcactgcccc tcctgtacat caggaaggcg actccttgag tcttggctct   14640
ggccgcctcc tccacctgca gtgagttaac tccttacct actctaggtc attgctcaaa   14700
```

                                    58

```
tgtcagcatc tcaatggggc cctccctgac taccctattt aaattctaca tactcccctt   14760
gaccccatgg acctcactca ccctattcca cttttattct tacaatttag cacttgttct   14820
cttctaacgt attctaagac ttactcattt attacattgt ttgccacccc ctctagtaca   14880
taaactccag aggggcaggg atttctgtct atttattcat ttcttatcc ctaggacata    14940
gaacagggca tagttcagag tattcaatgt tatcaatgaa tgaactagca gtagtaccag   15000
ttccagttag gcacagaatt aaatctaaat agaattaaat ctcatggtct gggttactca   15060
tggatagaaa attagatata atttaagaa gcctagaaag aaaaaattaa taatgtaaaa    15120
ataattataa tttgataata ataacaaaaa ctctgccagg cactgtggct caaatctgca   15180
atcccagcta ctcaggaggc tgaggtggaa ggatcacttg agaccagagt tcaagactca   15240
gcctaggcca cacggcaaga aactgtctct aaaaaaatta aaacttaaat ttttaaaaaa   15300
gaattctcaa agcgtcacaa aaactggaga ttaaggtaca ggaagtgtga agtaatatta   15360
ctatgctaat ggtttttttt ttttttagaa aggtataacc aaaagatttc tttctcaagt   15420
cgataaactg agaaagataa gcatatcttc caattaacag aggggaggga aaagcccagat  15480
acaacaaaat aagatataaa ttagtttcca gttgaaaaca agagtaggag ttattttgca   15540
tcacctcacc tgtgacctcc cccagcccaa aaaacactac tgataaacag ggtagaaaag   15600
catcatctca gataaagcag gaaaaactgc cacagtctca aaccacaaac tataagcaca   15660
cacctggcca accctgccaa gtctgggctc agtaggagga acgtgctgag agctaggatg   15720
taccaactta gacattctgt gggatacaga tgtccctgga agggtcacac catctcaaag   15780
gcacctgtaa tgcccactga ttacagccac catatgtgag agagaaactc agggcactta   15840
gagagtataa caagaacctt atgtcatctg agatgaagaa tcctcagccc tgcaaattaa   15900
ccaactcttt agaacaactg gcaaaacata aatatccaca acttttgttt cagtaattcc   15960
actcttagat atcaatccaa agtacatgag acagcagata cacacacaaa atggtatttta  16020
ctgcagcatt gtttataata gcaaaaaaca agaaataatc catatgtctc aataggataac  16080
tgggtacatg agggtatgta cccatcattc aaccatcaaa aagagtgata tggatgtcca   16140
cagatgaca taaaaagctg tgtgttacgt gaaaacaaac tcaagcagca gcaggatggg    16200
cttatgatag tcagtatgag ctaatttctg gaaaaaaaaa tctagtgtgt gcacagaaaa   16260
catctgaaag aacagaaaca aaactatcag cagaatattg agatgtttta ctaagttgta   16320
tatctatact gcttgtaatt tttaccccaa gcaagaatta ctttttggaa aaagaaaatt   16380
caggaaataa agcatttctt taaacttcat gtttaaacaa atggtgatgg aataaaagag   16440
ttcttattca tcataaacac acacagcaca catgcacgca tgtgcgtgag cacaccctt    16500
acttgataaa taccatgttg aatatttag tcttttccttt taggtttctat cccttcactc  16560
aaaatgcggt tataaataaa tgtactttc atgtgccttc tgcctaaacc cactttaata    16620
taactttaca gtcccattat cattatgtc tcaaagctag actcagcctg aaactacccct  16680
ttcatttgga acccttatta aaatgccaca tacagctcct tcaaataaaa acaaacccta   16740
ggacctgaca ctaggctcc tttgttgcta ctcataatgg ccaagttctg tgctataaat   16800
acatcttctt tcattttatt gctacatatc caagggtttt atatgttttt cttattatat   16860
cttaattcaa aacaccatca cgctcttttc cagatgaaaa taaggaaaag aaattgagca   16920
actgactgac ttaaagtcca taaaactata tagtagcaga gtcagcaaaa gaagaaacac   16980
acatctccca agtagaggct gaaaaccagt accattcacc tccagggtga gtatatataca  17040
gattacaaag tcacttctc taaatgttca aactgaatcc catcccata ctttaccact    17100
acctcgtaag aacagcctca gatcttgtta tagcctttt tttagcatgc tgaagccaat   17160
aaaatgcttc ccattcagca agagaaacaa gttctgaaac actgaataat ctgcccaggg   17220
cctatgaaca tttccactgt gagaaatgtt ccccactgtg tggagaagat ccttactctt   17280
ctccacacag gcagaaacatt agaaaaattc ttggatttcta tgatgcacag cttaggagtc  17340
tgttagcac aatttaagtc caaatagtta ttaaatcctc ctctgttcca gaaacagtgc    17400
taaatactgt gaatataaaa attgaaaaga tactctcctg gctcccaaga aagtcagcca   17460
gatagaggag acacagccac acaaatcact gtcacatgaa gctctacctc cctaacttca   17520
aacgaggggc taagtcacca agaatacagt agcaggttt actacgagta actactataa   17580
ttcaatactt tatcttccct tagaaaaactc ttctcccttg gaaatttatt tgcatttcta  17640
aataccattc cttactaaaa ggaagcaggg ctcctttgggg aaatagctga ttctatacca  17700
ggactatgaa atgaaaatgg tgagtctggg acatcccatg ttgcccagaa atcaaggaac   17760
tgcccaaaga ttaacagagt catgttaaat ggacctaaga gtgaaccaga aggagctcac   17820
tttgcccgc gtggaacaat ttcaagaaaa acatgacagt aatgaattat aaaacatgaa   17880
ttaaaataca tattggtact aaaaagagaa caaaaggatg tggctttgga taaagctctt   17940
cttcatggaa gaataccagc taataaatgt aaaggaaatg agagaattag aaaaattatc   18000
attttgtaaa ccttaatata ttcacctaga catgctaaaa ccactgagta aaaggctgct   18060
tgggaagagg atgctcacat gatctcagag tttcacacca cagataattt attagataca   18120
ggaaggaaga tgtgatcaag cttcctgtga cccccagcca ggcccacaa cactatgtgc   18180
ctccttgtga tgtgggagct acacagcatc gcccacacag cttctcgcca aaactgtttg   18240
aagctaatca caagggaga actggacagc ttctgaccat gagacgctcc accagacaac   18300
ttgcttggcc tctccaaaga aacttgcttg gcctctccaa agaaaactca gtttcattta   18360
aaaacaaaac taattattta aaaacaaacg aaaagcaagt tgtggacttg agctccaggg   18420
acagacgaga catactttc cctgttcttc ccagtaagtg gtaataaaaa ccctcaacac    18480
tagatataa acaaatataa gaaggttctg gaaggggaag aggaggcaga ctatccaggt    18540
gccttgaggc ccacagaaca acccagtgat gggttcactg ggtcttcttt ttgcttcatt   18600
atctcagact tggagctgaa gcagcaggca acttcaaaac accaaggggc acagattgaa   18660
aagccccaag aaaagcctgc cctctctagc caaaggacca ggaaggagac agtcctaatga  18720
```

```
gatggaacac atttagacag taactgccca tttaccagca ataactgagc agggagccta   18780
gacttccagt cttgtgagga cgtaccaagg tacccaacac ccccacaagg gctgagtaag   18840
gactgcgact tttatccctg catggcagta gtaaggagcc catcctcac ccgccagcag    18900
tgtcagggga acctggactt ccactcccac ccaggagtga tgaggccctc cctgctgagg   18960
tcatgtcaga ggaggcctag tggagattca gtgacttaac cttttcccag agataatgag   19020
gccaccttc ctccctcttc cccatggtg acagtgaaag cactgtggca agcagtaggc     19080
actcctaccc ctcctagcca gggaggtatc agggaggcca agtagggaac cagaataccc   19140
acaaccaccc agcagcaaca ggggtcccccc accccattgg gtgtcaatgg aagcagagcg   19200
gaaagcctgg atatttaccc ccatctagaa gtaacaagct gatgtccccc ttcttctact   19260
acaatggtgt tcaaaacagg tttaaataag gtctagagtc tgataacgta atacccaaat   19320
cgttgaagtt ttcattgagg atcatttata ccaagagtca ggaagatccc aaactgaaag   19380
agagaaaaga caattgacag acactagcac taagagagca cagatattag aactacctga   19440
aaggatgtta aagcacatat cataagcctc aacaggctgg gcgcggtggc tcacgcctgt   19500
aacccagcac ctttgggagg ccgaggcagg tggatcacaa gatcaggaga tcgagaccat   19560
cctggctaac acggtgaaac ccgtctctca ctaaaaatac aaaaaaaaat agcaaggcat   19620
ggtggtgggc acctgtagtc ccagctactc gggagcctga ggcaggagaa tggcatgaac   19680
ctgggaagag gagcagtgag ccgagatcgc accaccgcac tccagcctgg gcaacgagc    19740
aagacttcgt cccaaaaaaa aaaaaaaaaa aaaaaaagc ctcaacaaac aactacaaac    19800
gtgcttgaaa caatgaaaaa aaaatctg gcaaagaaat aaaagatata tattttggcc      19860
aggtgcagtg gctcacgcc tgtaatcccc gcacttttggg aggctgaggc aggcggatca    19920
cctggagtca ggagtttgag accagcctga ccaacatgga gaaacccgt ctctactaaa     19980
aatacaaaat tagccagtca tggtggcaca tgcctgtaat cctactacct caggaggccg   20040
aggcaggaga atcgcttgaa ctcaggaggt ggaggttgcg gtgagccgag atccgccat    20100
tgcacattgc actccagcct gggcaacaag agcaaaactc catctcaaaa aaatagatac    20160
atattttaat ggaaatttta gaattgaaaa atacagtaac caaattgaat ggaaagacaa    20220
catagaatgg aggggggcaga caaaataatc agtgaacttc aacagaaaat aatagaaatt    20280
acccaataaa aagaacagaa agaaaataga ctggccaaaa aataaagaag aaaaaagagg     20340
agcagcagga ggaatgatgg aaaaagagaa aggaaggaag gaaggaaggaa agggagggaa   20400
ggagtgaggg agaaagtctc aaagacctct gagactaaaa taaagatct aacacttgtc     20460
atcagggtcc aggaaagaga caaagatggc acagctggaa acgtattcaa aaaataatag     20520
ctgaaaactt cccaaatttg gcaagagaca taaacctata gattcgaaat gctgaacccc    20580
aaataaaaag cccaataaaa tccacaccaa aatacatcat agtcaaactt ctgaaaagac     20640
gaaaagaaa aacgtcttga aagcagtgag tgaaacaaca cttcatgtat aagggaaaaa      20700
caattcaagt aacagatttc ttacagaaat taaggaagcc agaaggaaat gacacaatgg     20760
ttttcaagtc ctgaaagaaa agaagtgtca acacaaaatt ctagattcag taaaaatatc     20820
cttcaagaat caatgggaaa tcaagacagt ctcagataaa gcaaaataag agaatatgtt      20880
gccagcagat ctcccctaaa ggaatggcaa aaggaagatc atgcaacaga ccaaaaaatg     20940
atgaaagaag gaatccagaa acatcaagaa gaaagaaata acatagtaag caaaaataca     21000
tgtaattaca ataaaatttc tatctcctct taagacttct aaattatatt gatggttgaa     21060
gcaaaaatta taaccctgtc tgaagtgctt ctactaaatg tatgcagaga attataaatg     21120
gggaaagtat aggtttctat acctcattga agtggtaaaa tgacaacact gtgaaaaatt      21180
acatacacac acacacgtaa gtatatataa atatatgtgt gtatatgtgt gtgtatatat     21240
atatatacat ataatgtaat acagcaacca ctaacaacac tatacaaaga catataacat      21300
aaaaacaatt tagataaatt gaaatggaat tctaaaaaat attcaaatac tctacaggaa    21360
gacaagacaa aaagagaaa aagagggagg acaaactaaa tttttttaaa acataaataa       21420
aatggtagac ttaagcccta acttatcaat aattacataa atgtaaatga tctaattata     21480
tcaattaaaa gacagagata gcagagtta tttaaaaaca tagctataag aaacctgctt       21540
tgggctgagt gcagtgactc acacttgtaa tcccagcact tcgggaggcc aggcaggagg     21600
gatcacctga ggtcaggagt tccagaccag cctggacaac atggtaaatac cccatctcta    21660
ctaaaaatac aaaaaaatta gccaggcatg gtggcacacg cctgtagtcc caactactca     21720
ggaggctgcg acacaagaac tgcttgaacc cgggcagcag aggtagcagt gggccaagat     21780
tgcgccactc cagcctgaac gacagagtga gactccacct cagttgaaaa acaaaaaaga     21840
aacctgcttt aaatatacca acatatgttg gttgaaatta aaagaataaa atatatcatg     21900
aaaacattaa tcaaaagaaa ggagtggcta tattaataac ataaaatga cttcagagaa      21960
aagaaaattt caagagacag gaataaaaag atcaagaaaa gatcctgaaa gaaaagcagg      22020
caaatcaatc attctgcttg gagattcaac accctctctt aacaactgat agaacaacta     22080
gacaaaaaaa tcagcatgga gttgagaaga acttaacacc actgaacaac aggatctaat     22140
agacatttac ggaacactct acccaacaat agcaaaataa acattctttt caagtattca     22200
ctgaacatat ccttagaccc tacctggggc cataaaacaa agctcactag tgattgccga      22260
aggcttggat ggacagtgga agagctgcat ggggaggggag aagtgacag ttaaagagtg      22320
taggatttct ttttgggata atgaaaatgt tccaaaattg attgtggtga tgttggcgca     22380
actctacaaa tataaaaaag gccattgaat tgtacgtttt aagtgggtga aacatgccat     22440
atgtggatta tatctaacgc ttttttaaaa cttaacacat ttcaaagaat agaagtcata    22500
cagagtgtgc tctactggaa tcaaactaga aagaggtaac tggaggataa cgagaaaaagc    22560
ctccaaatac ttgaaaactg cacagcacat ttctaaaatc atccgtgggt caaagatatt      22620
catttctgat attcattttt attgtttaat gtattttttaa aaatttctta agggaaataa     22680
actgactaaa aatgaatatg gctgggtgcg gtggctcacg cctgtgatcc cagcactttg     22740
```

60

```
ggaggccgag gctggtggat cacaagatca ggagttcgag accagcctgg ccaagatggt    22800
gaaaccccgt ctcaactaaa aaactacaaa aagtagccaa gcgcagtggc gggagcctgt    22860
ggtcccagct acttgggagg ctgaggtagg agaatcgctt gaacacaggc agcagaggtt    22920
gcagtgagcc aagattgtgc cactgcacgc cagcctgggc gacagagact gcctcaaaaa    22980
aaaaaaaaaa aaaaagaata tcaaaatttg tgggacatag ttaaagcaat gctgagaggg    23040
aaatttataa cactaaatgt ttacattaga aaagagaaaa agtttcaaat caatagtctc    23100
cactcccatc tcaagaacac agaagatgaa gagcaaaata aacccaaagc aagcaaaaga    23160
aagaaaatat aaaaataaat cagtaaaatt gaaaacagaa acacaataaa gaaaatcagt    23220
gaaacaaagt actgattctt cgaaagatta ataaaattga caaacctcta gcaaggctaa    23280
caaacaaaaa agaaagaaga cacggattac cagttattag aatgaaagca taattagaaa    23340
caactctaca cattataaat ttgacaatgt agatgaaatg gactaattac tgaaaaaaca    23400
caaattacca caactcaccc aatatgaaat agataatgg gatagcctga taactactga    23460
gaaaattgaa tttgtaattt taacactctt aaaacagaaa cattaaactt aatatttat    23520
aaatattaga taaggtaatt ataccccttcc ttaacaaata aaaacgacaa attattttgc    23580
agctaaagag atgtatgtac tgtgaaaaat atcttcagaa aaatagaact ttgtttgaag    23640
aataaggatt taaaaaatgt ttttaactct caagaagcaa atatctgggc ccagatggtt    23700
tcactgaaga attctaccaa atgtttaatg aagaattacc accaactcta catagcatct    23760
ttgagaaaac tgaagagaag ggaacatctc ccagttcatt ttatgaagtg ggtgttactc    23820
tgatactaga actgtataag gacagctact ctgacacac tgcctatggg tagctctgct    23880
ctgcaggaac agtcagaaaa aaaaaaaaaaa gaagcactgg acaagggcag tataaaaaaa    23940
gaaaactggg ccaggtgcag tggctcacac ctgtaatctc agcactttgg gaggctgacg    24000
ctggtggatc acctgaggtc aggagtttga gactagcctg gccaacatg taaaaccctg    24060
tctctactaa aatacaaaaa ttagccaggc agggtggtgg ggaaaataaa aaggaaaaaa    24120
aaacaaaaat aaactgcaga ccaatatcct tcatgagtat agacacaaaa ctccttaaac    24180
tccttaacaa aatattagca agtagaagca atatataaaa ataattatac accatgatca    24240
agtgggactt attccagaaa cgcaagtctg gttcaacatt tgaaaacaag gtaacccact    24300
atatgaacgt actaaagagg aaaactacat aatcacatca atcaatgcag aaaaaagcat    24360
ttgccaaaat ccaatatcca ttcatgatac tctaataaga aaaataagaa taaaggggaa    24420
attccttgac ttgataaagc ttacaaaaga ctacaaaagc ttacagctaa cctatactta    24480
atggtgaaaa actaaatgct ttccccctacg atcaggaaca aagcaaggat gttcactctc    24540
attgctctta tttaacatag ccctgaagtt ctaacttgtg caaaacgata agaaagggaa    24600
atgaaagacc tgcagattgg caaagaagaa ataaaactgt tcctgtttgc agatgacatg    24660
attgtctcat agaaaatgta aagcaactag gggtaggggg gcagtggaaa cacgctggtc    24720
aaaggatacc aaatttcagt taggaggagt aagttcaaga tacctattgc acaacatggt    24780
aactatactt aatatattgt attcttgaaa atactaaaag agtgggtgtt aagcgttctc    24840
accacaaaaa tgataactat gtgaagtaat gcatacgtta attagcacaa cgtatattac    24900
tccaaaacat catgttgtac atgataaata cacacaattt tatctgtcag tttaaaaaca    24960
catgattttg gccaggcaca gtggctcata cctgtaatcc cagcattta ggaggctgag    25020
gcgagcagaa aacttgaggt cgggagtttg agaccagaat ggtcaacata gtgaaatccc    25080
gtctccacta ataatacaaa aattagcagg atgtggtggc gtgcacctgt agacccagct    25140
acttgggagg ctgaggcacg agaattgctt gaacaaggga ggcagaggtt gcagtgagct    25200
gggtgccact gcattccagc ctggtgacag agtgagactc catctcaaaa aaaataaaat    25260
aaagcatgac ttttcttaaa tgcaaagcag ccaagcgcag tggctcatgc ctgtaatccc    25320
accactttgg gaggccgagg caggcagatc acaaggtcag gagtttgaga ccagcctgac    25380
caacatggtg aaaccccatc tctactaaaa aatatataaa ttagccaggc atggtgatcc    25440
tcagctactc aggaggctga ggcaggagaa tcacttgaac ccgggaggcag aggttgcagt    25500
gttgagccac cgcacctccag cctgggtgag acagagcaa ctccgtctcaa aaaaaaaaaag    25560
caaaataacc taatttttaa aacactaaaa ctactaagtg aattcagtaa gtcttttagga    25620
ttcaggatat atgatgaaca tacaaaaatc aattgagctg gacaaaggag gattgtttta    25680
ggtcagtagt ttgaggctgt aatgcacaat gattgtgcct gtgaatagct gctgtgctcc    25740
agcctgagca gcataatgag accacatctc tatttaaaaa aaaaaaaatt gtatctctat    25800
gtactgacaa taagcacatg ggtactaaaa ttaaaaacat aataaatact attttttaatt    25860
gcctgaaaaa aatgaaatac ttacatataa atctaacaaa atgtgcagga cttgtgtgct    25920
gaaaactaca aaacgctgat aaaagaaatc aaagaaagact taaatagcgt gaaatcttac    25980
atgcttatag gttggaaaac ttaatatagt aaagatgcca attttatcca aattattaca    26040
caggataaca ttattactac caaaatccca gaaaatttt acatagatat agagcaagatc    26100
atacaaaaat gtatacggaa atatgcaaag gaactagagt agctaaaaca aatttgaaaa    26160
agaaaaataa agtgggaaaa atcagtctat ccagtttcaa gacttacata gctacagtaa    26220
tcaagactgt gatattgaca gagggacagc tatagatcaa tgcaaccaaa tagagaacta    26280
agaaagaagc acacacaaat atgcccaaat gatttctgac aaaggtgtta aaacacttca    26340
acggggggag atatgtctct cattaaaggg tgtagagtca ttgcacatct ataggcaaaa    26400
agatgaacct gaacctcaca ccctacagaa aaattaactc aaaatgactc aaggactaaa    26460
cataagatat acatctataa aacatttaga aaaaggccac gcacggtggc tcacgctcgt    26520
aatcccagca ctttgggagg ccaaggcagg tggatcacct aaggtcagga gttcgagacc    26580
agccggatca acatggagaa ggcccatctc tactaaaaat acaaaattag ctgggacgtgg    26640
tggcacatgc ctgtaatccc agctacttgg gaggctgagg catgagaatc gcttgaaccc    26700
gggggggcaga ggttgcggtg agccaagatc acaccattgc actccagcct gggcaacaag    26760
```

                                         61

```
agcaaaactc caactcaaaa aaaaaaaaaa aaaggaaaaa tagaaaatct ttgggatgta   26820
aggcgaggta aagaattctt acacttgatg ccaaactaag atctataagg ccagtcggga   26880
tggctcatgc ctgtaattcc agcactttgg tcaactagat gaaaggtata tgggaattca   26940
ctgtattatt ctttcaactt ttctgtaggt ttgacattt tttagtaaaa aatgggacct   27000
aagacctgac gcagtggctc acacctgtaa tcccagcact ttgggaggcc ggggcaggtg   27060
gatcacacgg tcaggagttc gagaccagcc tggccaacat ggtgaaaccc cgtctctacc   27120
aaaaatataa aaaattagcc gggtgtcatg gtgcatgcct gtaatcccag ctactgagga   27180
ggctgaggca ggagaatcac ttgaacctgg gaggtggaag ttgcagtgag ccgagattgt   27240
gccactgcac tccagccttg ggtgacagag cgagactccg tctcaaaaga aaaaaaaaaa   27300
aaagaatatc aaacgcttac tttagaaact atttaaagga gccagaattt aattgtatta   27360
gtatttagag caattttta gctccatggc attgttaaat agagcaacca gctaacaatt   27420
agtggagttc aacagctgtt aaatttgcta actgtttagg aagagagccc tatcaatatc   27480
actgtcattt gaggctgaca ataagcacac ccaaagctgt acctccttga ggagcaacat   27540
aaggggttta accctgttag ggtgttaatg gtttggatat ggtttgtttg gccccaccga   27600
gtctcatgtt gaaatttgtt ccccagtact ggaggtgggg ccttattgga aggtgtctga   27660
gtcatggggg tggcatatcc ctcctgaatg gtttggtgcc attcttgcag gaatgagtga   27720
gttcttactc ttagttccca caacaactgg ttattaaaaa cagcctggca ctttcccccа   27780
tctctcgctt cctctctcac catgtgatct cactggttcc ccttcccttt atgcaatgag   27840
tggaagcagc ctgaagccct cgccagaagc agataagtgat gccatgcttc ttgtacagcc   27900
tacaaaacta tgagccaaat aaacctttt tctttataaa ttatccagcc tcaggtattc   27960
cttatagca agacaaatga accaagacag ggggaaatca acttcattaa aataatctat   28020
gcagtcacta aacaaataag aacaagaggc tccagaagtg ggaagccaat acccagagtt   28080
cctacaatac agtatctgaa aagtccagtt tccaaccaaa aaatatatat aacaggccg   28140
gacatggtag cttatgtctg taatcccagc actttgggat gctgaggcgg gcagatcacc   28200
ctaggtcagg agttcgagac cagcctggcc aatatggcaa aaccccgtct ctactaaaaa   28260
tacaaaaatt agccaggcat ggtggtggat gcctgtaatc ccagctactc gggaggctga   28320
ggcagggaat cacttgaacc caggaggcag aggttgcagt gagccgagat cacgccactg   28380
aactccagcc tgggcaacaa agtgagactc cacctcaaaa aaaaaaaaaa tatacatata   28440
tatatgtgtg tgtgtgtgtg tgcgcgcgtg tgtgtatata cacatacaca tatatacata   28500
tatacagaca cacatatata tatgaagcat gaaaagaaac aaggaagtat gaaccatact   28560
ttctgtgtt atgataggat ggggtatcac ggggaaagta gacaagggaa actgcaagtg   28620
agagcaaaca gttatcagat ttaacagaaa aagctttgg agtaaccatt taaatatgt   28680
ccacagaatt aaagaaaagc gtgattaaaa aaggaaagga aagtatcata acaaattac   28740
tccaaaataga gaatatcaat aaaggcatag aaatttataaa atataataca atggaaattc   28800
cggagttgaa aggtagaata actaaaattt aaaattcact agagaaggtt caacactata   28860
tttgaactgg cagaagaaaa atttagtgag acaaatatac ttcaatagac attattcaaa   28920
tgaaaaataa aaagaaaaaaa gaatgaagaa aaataaacag aatctcagca aaatgtggca   28980
caccattaat cacattaaca tatgcatact gagagtaccg gaagcagatg agaaagagga   29040
agaaaaaata ttcaaatgat ggccagtaac ttcctagatt tttgtttaaa agcaataacc   29100
tatacaatca agaaactcaa tgaattccaa gtaggataaa tacaaaaaga accacaaaca   29160
gatacaccat ggtaaaaatg ctgtaagtca aaaacagaga aaatattgaa agcagcctga   29220
ggaaaactta taagagaacc tcacttacaa agaaacatca cttataaaag aaccacaata   29280
atagaaacag ttgacctctc atcagaaaca atgaatgata acatattaga agtgctcaaa   29340
gaaaaaaaaat aaagatttcct atatacgaca aagctgtctt tcaaaaaatat acatccaaaa   29400
ggattgaaac caggttcttg aagagttatt tgtacatcca tgttcatagc agcattattc   29460
acaatagcca aaaggtagaa gcaacccaag ggttccatcga caaataaaata aaatgtggta   29520
tatgtattaca caatggaatt tattcagtat taaaaaggaa tgaaattctg acacatgcta   29580
caacatggct aaaccttgaa aacactatgc taagtgaaat aagccagcca caaaaggaca   29640
aataccatat tacttcactt gtatgaaata cctagggtag tcaaattcag agatagaaag   29700
taaaacagta gttgccaagg gctgagggag ggagtaacgt ggagttattg ttgaatgggt   29760
acagatttc agtttgcaa gataaaaaga gttctggaga cagatggtgg tgagggtggt   29820
acaacaatac aaatatactt tatactactg aacagtatac ttaaaaatga ttaaaggaca   29880
gaaaccccgt ctctactaaa aatacaaaaa aattagctag gtgtggtggc gggcacctgt   29940
aatccagct acttgggagg ctgaggcagc agaattgctt gaaaccagaa gcgcggaggtt   30000
gcagtgagtt gagattgcgc caccgcactc tagcctgggc aataagagca aaactccgtc   30060
tcaaaaataa aaaataaaa aaaatttaaa aatgattaag caggaggccg gacgcggtgg   30120
ctcacaccta taatgccagc actttgggag gccgaggcag gcgatcactt gagacagga   30180
gtttgagacc agcctggcca acatggcaaa accctgtctc tgctaaaaat aaaaaaatta   30240
gccaggcatg gtggcatata cttataatcc cagctactgg tgagactgag acacgagaat   30300
tgcttgaacc caggaggcag agattgcagt gagtcgagat cgcgccactg cactccactc   30360
tgggcgacag agcaagattc tgtctcgaaa aaacaaaaac aaaaacaaaa agcaaaacca   30420
aaaaataatt aagcaggaaa cgagattgct gctgagggagg agaaagatgt gcaggaccaa   30480
ggctcatgag agcacaaaac ttttcaaaaa atgtttaatg attaaaatgg taaattttat   30540
atgtatctta ccacaaaaaa aagggctggg gggcaggaaa tgaaggtgaa ataaagacat   30600
cccagagaaa caaaagtaga gaatttgttg ccttagaaga aacaccacag gaagttcttc   30660
aggctgaaaa caagtgaccc cagagggtaa tctgaattct cacagaaaat tgaagcatag   30720
cagtaaaggt tattctgtaa ctatgacact caaatgcat attttttcct ttcttctctg   30780
```

                                                                          62

```
aaaatgattta aaaagcaatt gcataaaata ttatatataa agcctattgt tgaacctata    30840
acatatatag aaatatactt gtaatatatt tgcaaataac tgcacaaaag agagttggaa    30900
caaagctgtt actaggctaa agaaattact acagatagta aagtaatata acagggaact    30960
taaaaataaa attttaaaaa atttaaaaat aataattaca acaataatat ggttgggttt    31020
gtaatattaa tagacataat acaaaaatac cacaaaaagg gaagaagaca atagagactac   31080
ataggaataa cattttggta tctaactaga attaaattat aaatatgaag tatattctgg    31140
taagttaaga cacacatgtt aaaccctaga tactaaaaag taactcacat aaatacagta    31200
aaaaaataaa taaaataatt aaaatgtttg tattagtttc ctcagggtac agtaacaaat    31260
taccacaaat tgagtggctt aacacaactt aaatgtattt tctcccagtt ctggaggcta    31320
aacacctgca atcaaggtga gtacagggco atgctccctg tgaaggctct aggaaagaat   31380
cctcccttgt ctcttccagc ttccagtggt tctcagtaac cctaagtgct ccttggcttg   31440
tagctatatc attcctagca accagaaaga agaaaataat aaagattatg gcaaaaaata   31500
atgaaatcaa aaggagaaaa atggaaaaaa ataaataaaa ccaaaagcta gttctttgaa    31560
aagatcaacc aagttaacaa acctttaac tagactgaca aaaaggaggt agaaatctaa     31620
ttactagaat cagaaataaa agagggggaca ttactaatga gggattagaa aagaatacta   31680
cgaacaaatg tgtgccaaca aattagaaaa cttagtgaa atggacaggt tcctaggaca    31740
acatcaacta ccaaaattta ctcaagaaga aagagacaat ttgaatgagc tataacaagg   31800
gaagagactg aattgacaac caagaaacta tccacaaaga aaatcccagg cccagaagat   31860
ttcactgtga aattcttctca aacttataaa tataaattaa catcagttct tcacaaactc   31920
ctccaaaaaa aagaacagat ctctatttac aggcgatacg atctttagaa aatcctaagg   31980
gaactactaa gacactatga taactgataa acaagttcag caaggctgca ggatagaaaa   32040
ccaatataca aaaatctatt atattttctat acacttgcag tgaacaaccc aaaaatgaga   32100
ttaagaaaat aattcaattt acaataacat caaaaagaat aaaaacactc aaaaataaat   32160
ttattcaagt aagtgcaaaa cttatactct agaagctaca aaacactgtt aaaagaaatt   32220
aaaggtttac ataaatgaaa aactatccca tgttcatgga tcaaaagact tattactggc    32280
aatgctctcc aaattgatct ataaattcaa caaaatcctt atcaaaatcc cagatgaggc    32340
tggggtggc ggttcatgcc tgtaatccca gcactttggg aggctgaggc acgcagatta    32400
cctgaggtcg ggagctcgag atcagcctga ccaacatgga gaaaccctat ctcttctaaa    32460
aatacaaaat tagtcaggcg tggtggcaca tgcctataat cccagctact cgggaagctg    32520
aggcaggaga atcgcttgaa cccaggaggc agaggttgca gtgagccaag atcgtgccat    32580
tgcactccag cctgggcaac aagagcaaaa ttccatctca aaaaaaaaaa aaaaaaaatc    32640
ccagatgact tcactgttga aattgaaaag attattctaa aattcacatg gaattgcaag    32700
accttgagaa tagccaaaac aaacttgaaa aacacgaaca aaatatagga tgactcactt   32760
gccaattgca aatgttacga cacagcaaca gtaatcaaga ctgtgtggta ctggcaaaaa    32820
acacatacat acatacatat caatggaata taattgagag tacagaaaca agcctaaaca    32880
tctatggtaa gtgctttttct attttttttct ttttttttttt cttttttgta gagatagaat  32940
ctcaccatgt tgcccaggct ggtcttcaac ttctgggctc aagcaatcct cccactgtgg     33000
cctcccaaag tgctgggata actggcatga gccaccacat ccagcccaga tgattttcaa     33060
aaaagtcaac aagaccattc ttttcaacaa ataggtctgg gatgatcaga tagtcacatg    33120
aaaaaaaaaa tgaagttgga ccctccatca cactaaagtg ctgcgattat aggcatcagc    33180
caccacatcc agcccaaatg attttcaaaa aggtcaacaa gaccattctt ttcaacaaat    33240
aggtctgga taatcagata gtcacatgaa aaaaaaaatg aagttggacc ctccatcaca    33300
ccatatgcaa aaatttaattc aaaaatgaat tgatgactta aacgtaagag ttacgactgt    33360
aaaactctta gaaggaaaca tacgggtaaa tcttaaagac gttaggtttg acaaagaatt    33420
cttagacatg acaccaaaag catgaccaac taaggtaaaa tagggtaaat tgtacctacc    33480
aaaatgaaaa aacttttgtgc tggaaagac accatcaaga aatggaaagc caaaatagcc    33540
aaggcaatat taagcaaaaa gaacaaagct ggaggcatca tactacctga cttcaaagca    33600
acagtaacca aaacagcatg gtactagtag aaaaacagac acatagacca atggaacaga    33660
ataaagaacc caaaaataaa tccacatatt tatagtcaac tgatttttga caatgacacc    33720
ccttcaataa atgatactag gaaaactgga tatcgatatg cagaagaata aaactagacc    33780
cctatctctc accatataga aaaatcaact cagactgaat taaagacttg aatgtaagac   33840
ccaaaactat aaaactactg gtagaaaaca taaggaaaaa cgcttcagga cattggtccca   33900
ggcaaagatc ttatggctaa aacctcaaaa acacaggcaa caaaaacaaa aatggaaaaa    33960
tagcacttta ttaaactaaa aagctcctgc acagcaaagg aaacaacaga atgaaaagac    34020
aacctgtaga atgggagaaa atatttgcaa actatccatc catcaaggga ctagtatcca    34080
gaacacacaa gtgactaaaa caactcaaca gcaaaaaagc aaataatctg gtttttatat    34140
gggcaaaaga tctgaataaa cattctcaaa ggaagacata caaatgtcac tatcattctg    34200
ccagtaccac actgtcttga ttacttgtta gtgtataaat ttttaaattg gaaggatgga    34260
gtcatcctac actttgttct tgttttttcaa gtttgtttttg gctattctgg gagccttgca   34320
agtataaaat agccaacaag tatgaaaaaa tgctcaccat cactaatcat caagaaagatt   34380
aaaatcaaga ccactatgag atatcctctc actccagtta gaatggctac tatcaaaaag   34440
acaaaatata atggatgcgg gcaaagattt ggagaaaggg gaactcctat acatgtttta    34500
tagggatgca aattggtaat ggccattatg gaaaataata ctgaggtttt tcaaaaaact    34560
gaaaatagaa ctaccatatg atccagcaac cctactactg ggtatttatc caaaggaaaag   34620
aagtcagtat actgaagaaa tatatgcact ttcatgttaa ttgcaaacat gttcacaaca    34680
gccaagacag ggaataaatc taaatgtgca tcaacagatg aatggataaa gaaaatgtgg    34740
catatacact caatagaata ctattcagcc attaaagaag aatgaaaatcc tgtcatccca   34800
```

```
gcaacatgga tgaacctgga ggacattata tttaatgaaa taagtaaagc acaaaaagat   34860
aaacagtaca tgttctcact cagacatggg tgctaaaaag aaaatggggt cacagaatta   34920
gaagggggag cttgggaaaa gttaatggat aaaaatttac agctatgtaa gaagaataag   34980
ttttagtgtt ctatagaact gtagggcgag tatagttacc aataacttat tgtacatgtt   35040
caaaaagcta gaagagattt tggatgttcc cagcacaaag gaatgataaa tgtttgtgat   35100
gatggatatc ctaattaccc tgattcaatc attacacatt gcatacatgt atcaaattat   35160
cactctgtac ctcataaata tgtataatta ttacgtcaac aaaaaaagga aaaaaaagaa   35220
aattaagaca acccacataa tggaagaaat aaaatatctg caaattatat atatctgata   35280
aatatttaat atttataata tataaagaac tcctacaact caagaacaac aacaaaacaa   35340
cccaattaaa aaatgggtaa aagccttgaa tatacactta tctaaagact ataatcaaag   35400
ggccaataaa gacacgaaaa gatgctcaac atcactagtc atcagggaaa tataaatcaa   35460
aaccacaatg tagaatgtag acaccacttc atatgcacta ggatggctag aataaaaagg   35520
taataacaaa tgttggtaag gatgtgaaaa aatcagaaac ctcattcgct gctgttggga   35580
atgtaaagtg atgcagccac tttggaaaac ggtctggcag ctcctcaaat tattaaatac   35640
agagttaccg tatgacccag gaatattcct cctgggtcta taaccaaaaa aatgaaaaca   35700
tatatccaca taaaaactg tacatgggca tttatagcaa cattattcat aacagcaaaa   35760
gtggtaagaa cccatatgcc catcatctga tgaacaggta aataacatgc ggtattatcc   35820
atacactaga atattatctg cccatacaag gagtgacatc cagctacatg ctacaaggat   35880
gaatctcgga aaccttatgc taagtgaaag aagccagtca caaatgacca cagattatga   35940
tttccatgcat cggaaatgac cagaataggg aaatctatag agacagaaag tagattaagtg   36000
gttgggtggg gctgggagga caggtagtac actacttcc cagaactact cagacaaagt   36060
accacaaact ggggagctta aacatagaaa ttgatttcct cacagttctg gagactagga   36120
ctctgagatc aaggtgtcag cagagctggt tctttctgag ggccctgagg caaggctctg   36180
tcccaggcct ctctccttgg ctggcaggtg gccatcttct ccctgcgtct tcacatcatc   36240
ttttctctgt gtgtgcccat gtccaaattt tgattggctc attctgggtc atggccaatt   36300
gctatgcaca aagtgaagtc tacttccaaa agaagggaag agggaacact gactaggcta   36360
aacttatagt cattttaatg tccgctttc ctatgagatt gtgaacacac agaagtaggg   36420
ttttatctta cattgtgcaa agtttaataa gaaaaataga attcaagaga agcagttcaa   36480
tagcaggaat ttaatatggg aactaattac aaggtttag gcaggactaa aaagccagtt   36540
gggatggtga gccaacccag agattagcaa cagtgggacc ccatctacct accaccactg   36600
aagctggaag gataaaggag gggctattat cagagtccac aagccagtgt cagagtcctt   36660
ggctggagct gggaccaccc tagagacact gtgcaaagca gaaaacaagg gggaaaaacc   36720
ctgacttctc ccttcctccc accttttcaat ctcccactag tgcttcctac tagccatact   36780
tggccagaga cagtgacaag gaacactgca aaatgaagtt tgtaggaatc atctccctct   36840
gagacagaga aatatggaag ggtagaaaat gaatcagagg ataaagagaa aaaaaccctga   36900
gtactatctt atttatcttt gtatctccag tgcctaatct gtctctcaaa aaaggaaagc   36960
aattgagaga aactgaaaac tccaattgaa atgaaagaat ggagaattac tggactagaa   37020
gagaagagaa aaatttattc cgcatagagt aaacaagaat ggattcacaa aggacgtgaa   37080
gaatgaaaag ctataatcag caaagatttg ccagagaaat taaaaagtgg taaactcagc   37140
cacgctgtac aacctgaagg cacaatgcat gaaaacgttt caagaaatga caagatttga   37200
agtcaaattc taagtgcttt tccagaatct ctcaagacga ttatatagct accccattt   37260
attaaataaa atgggaaactt actaaacttt ccccttgtat taaactaaca tatgtcctaa   37320
tagcaaacga ttctggaatt cctagagtaa aatatattc gtcaaagtgt acacgtaaat   37380
taatattctg ctgacctcct tttgctattt aggatatttg tatacacatc acacgtaaat   37440
ttggtctata gtttacatct acgggcttat actgttctttt ttttcatttt tttaaaattt   37500
ccaacccca gtatccatat actgctctct atcagggtta ttttaacttt gtaaaatcag   37560
ctgagatgct ttccatgttt ttttttttta ttttctgcca catttgaata gcataggagt   37620
taccaccatc aacctttggat tatttaagca ttcacgattc cacgtgtgga tttttttattc   37680
agagtctttc ttgtcattcc tgctatcagc acagaaccca atctcagctt tccagctata   37740
ctctcaccc atggaattttg cagatgaagt tcaaaaggac ctttgcatta tcctgcctcg   37800
ccctcttccc ccttcattta gacatcacct tcttctagaa cgtcttacct gacatgccct   37860
gctcccaacc cctgctgccc aattgtgtgc tctcccgtgt cctggcctgc catcctcttt   37920
agtaattgcc tgctccctca tctgtctccc caccccagaca ttaagctgaa tagactggat   37980
ttgtgtcttg tccatcacta taatctcagc acctagtacc tagtaggtac ttaccatgta   38040
ttcattagca aaatgttatg tataaccttg caccttaaaa acaagagaag gaagacaaaa   38100
ttaagtctta agactatggt ttagaacatg gatcagaaac tacagtctgc agcccaaatc   38160
cagaccaaat gaagagacca tgttcattta catacaacct atagcagctt tcacactaca   38220
ggagcagagc taagtagttc caagggaaca cacggccctg caaagcctaa aatatttact   38280
ctatagctct tcacagaaaa agttttcaga tccctcgttt agaactcttg ttcatatgca   38340
atttcactaa accatagtttt tttgggtttg tttggttttt tttgtgcaaaa aggaatgagc   38400
cgatccagaa aaggttgaaa agaatgaatc attactgctg aaagaatgtg cacacagtcc   38460
gtcagtattc tgctgccatg ctgacaccca tccaatagtg tcatgagatg cagcagctac   38520
tactgtgttc tcaatgccga gtccacccac tccataacca tgtccaagca atcttgggaa   38580
catcatcacc atgcttgttt atccttaagg tattgcctca catacagcag tggctggtca   38640
taaagtcaaa tgacactagt ggccaggagg tcaagagaat gagtgaggac aggtgggtag   38700
gcagcccagg ccctagcaac agcaggagct caccctcag tcactctagc caggactgaa   38760
atacttttca cccttcaag agagactagg aatctggatt tttatgtgaa atatcttgat   38820
```

                                    64

```
tactaaatgt tgtcaacaga catgtcaaaa ggtaaaacta agtaagttca tggggcagat    38880
tgactattca ggttatagaa ttaaggattc ttatccaaca cagataccaa ccaaaagct     38940
gacgtataac atattaggag aaactatgtg cactgtcgaa acatcaacaa ggggctaatg    39000
tctaaaatag tctatattgg attccagttg aaacatgggg aaaggacatg aacaggcaac    39060
ttatgtcaat ggaaactcaa aaagataaca agcatatata aaagcattct caaattcagt    39120
agtaaacaga cagatgcaaa taaaaagagg gaaactgctg ccgggcacag tggctcacac    39180
ctgtaatccc agcactttgg gaggccgagg cgggcggatc atgaagtcag gagatcgaga    39240
ccatcctggc taacatggtg aaaccccgtc tctactgaaa acacaaaaaa ttagccaggc    39300
gtagtggtgg gcaccagtag tcccagctac tcaggaggtt gaggcaggag aatggcatga    39360
acccaggagg cggagattgc agtgagccga gaccatgcca ctgcactcca gcctgggcga    39420
ctgagtgaaa ctccatctca aaaaataaa taataattat aattataata ataataaata    39480
gtaaataaat aaaaagagag agactgctaa agtctagaaa gttgaatgat gccaagcgca   39540
tgcaaagatc agggccttgg gatggccggg tgcagtggct cacgcctgta atcccaccac    39600
tttgggaggc caaggcgggc ggatcatgag gtcaagagat caagaccatc ctggccgaca    39660
cagtgaaacc cggtctctac taaaagtaca aaaaaatata tatatatata tatattatta    39720
tattatatat atatatatca gagccttggg aatccttgtg tgctgctggg gaaggtagtg    39780
gtgcagccac ccttgacagc aatctggcag tacttggtta tattaagtat aggcacacac    39840
cacgaccagg cagtcctact cctgggtcta aatcccaaag aattctcaca caagtccata    39900
aggagacatg tacgaggctc attcagcatt actgggagtg ggaatcaacc tgggtgtcca    39960
tctacagggag acgagatgga caaaatgtgg tggatattaa gaccagaatc accaagtaac   40020
agagatgggt ggtgagtgac aatcctaaga tacagaataa aggctagaac atgatgccat    40080
tcatgtaaat taaaaataga tgcacacaaa gcagtatacg cgtgaccctt gaatagcaca    40140
ggtttgaact gcctgtgtcc acttacatgt ggattttctt ccacttctgc taccccaag    40200
acagcaagac caacccctct tcttcctcct cccctcagc ctactcaaca tgaagatgac    40260
aaggatgaag acttttatga taatccaatt ccaaggaact aatgaaaagt atattttctc    40320
ttccttatga ttttctttat ctctagcttа cattattcta agaatatggt acataataca    40380
catcacacgc aaaataaatg ttaattgact gtttatatta tgggtaaggc ttccactcaa    40440
cagtaggctg tcagtagtta agttttggga gtcaaaagtt atacacagat tttcaactgt    40500
gcaggcaatc agttccctg accccctcat tgttcacggg tcaactgtat atacacaaaa    40560
gtattatatg aacctcatta gaatagctgt ctatagggag aagagaatga gagtgggata    40620
aaacggaatg aacaaataaa ccaacaaatg cattaacaag caaaacaaca gaggggcttg   40680
catgggccag tgatgataaa gggctaagaa tgagaatata attaattcaa ttcctcacac    40740
ctgaggtcta aaaccaagga aagggagggc caggcgtgga ggctcacgcc tgtaatccca    40800
gcacttttggg aggctgaggc gggcggatca caagtttagg agtttgagat cagcctggcc    40860
aacacagtga aagcccatct ctacaaaaaa tacaagaatt acccaggtgt ggtggcacat    40920
gcctgtagtt agctactctg gaggctgagg caggagaatc acttgaaccc aggaggcgga    40980
ggttgcaggg agccgagatc acaccattgc actccagcct gggtgacaga gtaagacttt    41040
gtctcaaaaa aataaaaaaa ataaaaaaac agagaaaggg aggaaactag atccaggctg    41100
actagataca gcctttagag ttagaaaaga tgatttgaca atctaagccc acactcagat    41160
tgaatgaaat tgaaaagct ttcaaactaa aacatttaat tacaccatct gctgcagaca    41220
gaactcagac aactcaaaca ggtaatgtca gcgtggtgtt ttatatcacc accttcaaca    41280
cagaataaaa atcagctgca tgtgaagcag tgactagaat gaagaaaagg ctgcttctta    41340
cttccttcta gtggttcttt ccgaaaacat taataggcac cagctctatg catgtcaccc    41400
tgcagggaga catggggtat ataactatga cttactgttc attcctcaag gaattcccaa    41460
tcttgtggaa gattatacac aatgaggcaa caaaaactat ccaataaaac cacggaaaag    41520
aagccagtga caagaaagcc agtgatgaaa ggccctgtga gcagagctga tggccatttg    41580
gggaagaaag accaacatgg atgggggtga tcagggtggc tccgtgggaa agctggaaga    41640
gaagtggcag atctctgagc tggatgatgg gccactacca tctgtatatg gctaattaaa    41700
gaccatgtgt ggatttttta ttcagctctt tcgtgtcatt cctgctatca gcacagaacc    41760
caatctcaac tttccagcta tattgagcta aacttctcac ctcatggaat ttgcaagaca    41820
agttcaaaag gatccttgcc ttttcaaaat aattttgaat ggttgagtag tccctctgtg    41880
ctctctcact gacaccctct caaggctgct gagcacgtgc catgctatgg cttttctccaa    41940
catcaggaaa tgttctccac tcagtttcac cttaatacaa atgtgttctc tcttcagaga    42000
aggcaaaaaa attcatgacc atctgactgg gagaagtcat ttctaggtaa agtgtccatc    42060
tttttctgag gaacacagga ggaaaatctt acagaaaaga gttaacacag caggcctaag    42120
actgcttttt aaaataaata aataaataaa taaataataa aataaata a ataaataaata    42180
aataaatgaa tgataggggtc ttctgtattg gccaggctag tctcaaaattc ctggcttcaa    42240
gagatcctcc caccttggtc tcccacagtg ttgggattaa agacatgagc cattgtgctt    42300
ggcccaagac tgttatctt aaaaagtctc ataaaaagca tggttaatcc ttggctggca    42360
cctgggaact tagatttcag aagggttccc accatccaac ctggaaagag ggactcactg   42420
tgcctaaatt attgtgttgt ttatgctgaa ctcctgcttt tcttcaggta gcgtggaatg    42480
tggtatgctgc tgggcaaagg gggcctgcat gaccagcccc caataaaaac cctgggtgct    42540
gggtctctag tgagtttccc tggtagacag catttcacat gcgttgtcac agctccttcc    42600
tcggggagtt aagcacatac atcctgtgtg actgcactgg gagaggatgc ttggaagctt    42660
gtgcctgact tccttttggac ttgcccccat gcacctttcc ctttgctgat tgtgctttgt    42720
atccttttcac tgtaataaat tacagccgtg agtacaccac atgctgagtc ttccaagtga    42780
accaccagat ctgagcatgg tcctgggggc cccaacaca gaaatataatt ataaagagacc    42840
```

                                      65

```
aaggactggg catggtggcc catgccggta atctcagcgc tttgggaggc cgaggcagga  42900
ggaccagtta agcccaaaag ttcaaagtta cagtgaccta tgactgcgcc aatgcactct  42960
aacctgggag acagagcaag accctgtccc caaaacaata aactaaacac atacttctgc  43020
cttccaagtg tcttaaaatt caatggaatg gtagaaacat ttttaaaaca ctaaatcaaa  43080
agaaacctgg aaaacaagag tgccgatggc caactaaaat gtctaggaaa tttctgaaaa  43140
gtaaaaagta ctcagaacca gattacctga gcaaaccata gcccaataca agcttgggag  43200
gaggctgtta tgcagaagga aatggtaaca ggtttccagg aacagacttg taacagcaga  43260
tagaacgaca gaggtagaac ctgacaaggt gattacctgg ggaactgcag tctgaatgac  43320
caggactgtt ggacccttcc cctcacatgg aatacacacg ccactcagca gcacaccaca  43380
gctcttcaac aatcacagga ggcacgctac gcctagtaag acaggaaaaa aggaattctc  43440
aaacttcgaa gatgaacaca taaagaatca ccaagttttt attcagtatg atgaaacagg  43500
gacactgaat caacagaaca caaacccaag caaagataat tactagagca catagaagaa  43560
attattagat aattcttggga agacctaagg ggacattata aagagcaagc agttggtatg  43620
tgacgatctt tgtgatatac caagaaataa aaacacagga tgaagaccag atagagaata  43680
atgctactat ttgtgcaaaa aaggagaaat ggagaatctg attcatattt gcttgtattt  43740
gcatgaagaa actttggaag gtacataagt aactaacaac aatggttacc tacttgtaag  43800
gcgagagaag taagaggaca ggaatggtgg gaacaccttt tgtgtccgga attggtgggt  43860
tcttggtctg acttggagaa tgaagccgtg gaccctcgcg gtgagcgtaa cagttcttaa  43920
agggggtgtg tctggagttt gttccttctg atgtttggat gtgttcggag tttcttcctt  43980
ctggtgggtt cgtagtctcg ctgactcagg agtgaagctg cagaccttcg cggcgagtgt  44040
tacagctctt aagggggcgc atctagagtt gttcgttcct cctggtgagt tctgtgactg  44100
gctagcttca ggagtgaagc tgcagacctt cgaggtgtgt gttgcagctc atatagacag  44160
tgcagacaca aagagtgagc agtaataaga acgcattcca aacatcaaaa ggacaaacct  44220
tcagcagcgc ggaatgcgac cgcagcacgt taccactctt ggctcgggca gcctgctttt  44280
attctcttat ctggccacac ccatatcctg ctgattggtc catttacag agggtggctg  44340
gctccatttt acagagaacc gattggtcca tttttcagag agctgattgg tccattttga  44400
cagagtgctg attggtgcgt ttacaatccc tgagctagac acagggtgct gactggtgta  44460
tttacaatcc cttagctaga cataaaggtt ctcaagtccc caccagactc aggagcccag  44520
ctggcttcac ccagtggatc cggcatcagt gccacaggtg gagctgcctg caggatccgc  44580
gccctgcgcc cgcactcctc agccctctgg tggtcgatgg gactgggcgc cgtggagcag  44640
ggggtggtgc tgtcagggag gctcgggccg cacaggagcc caggaggtgg gggtggctca  44700
ggcatggcag gccgcaggtc atgagcgctg ccccgcaggg aggcagctaa ggcccagcga  44760
gaaatcgggc acagcagctg ctggcccagg tgctaagccc ctcactgcct ggggccgttg  44820
gggccggctg gccggccgct cccagtgcgg ggcccgccaa gcccacgccc accgggaact  44880
cacgctggcc cgcaagcacc gcgtacagcc cgggttcccg cccgcgcctc tccctccaca  44940
cctccctgcca aagctgaggg agctggctcc agccttggcc agcccagaaa ggggctccca  45000
cagtgcaggg gtgggctgaa gggctcctca agcgcgggca gagtgggcac taaggactga  45060
gaggcaccga gagcgagcga ggactgccag cacgctgtca cctctcactt tcatttatgc  45120
ctttttaata cagtctggtt ttgaacactg attatcttac ctatttttttt tttttttttt  45180
tgagatggag tcgctctctg tcgcccagac tggagtgcag tggtgccatc ctggctcact  45240
gcaagctccg cctcccgggt tcacaccatt ctcctgcctc aacctcctga gtagctggga  45300
ctacaggcaa tcgccaccac gcccagctaa tttttttattt tatttttttt ttagtagaag  45360
cggagtttca ccatgttagc cagatggtct caatctcctg aacctcgtgat ccatccgcct  45420
cggcctccca aagtgctggg attacaggcg tgagccactg cgcccttgcct atcttaccta  45480
tttcaaaagt taaacttaa gaagtagaaa cccgtggcca ggcgtggtgg ctcacgcctg  45540
taacccagca actttgggag gccgaggcgg gcggatcacg aggtcaggag atcgagatca  45600
tcctggttaa cacagtgaaa ccccgtcgct actaaaaata caaaaaatta gccgggcgtg  45660
gtggtgggca ccggcagtcc tcgctactcgg ggaggctgag gcaggagaat ggcgtgaacc  45720
tgggaggcag agcttgcagt gagccgagat agtgccattg ccttccagcc tgggcgacag  45780
agcgagactc cacctcaaaa aaaaaaaaaa aaaatagaga cccggaaagt taaaatatg  45840
ataatcaata tttaaaaaca ctcaagagat gggctaaaga gttgacggaa caaatctaaa  45900
tattagattg gtgacctgca aaaccagccc aaggaacatc ccagaatgca cccccataaag  45960
ataaagagag catttccgct gggcacagtg gtatggcagg ggaattgcct gagtccaaga  46020
gttgcaggtc acattgaacc acaccattgc actccaggcc tgggcaacac agcaataatt  46080
tgtctcaaaa aaaaaaaaaa ttaaattaaa aaagacagaa tatttgagag aaaaaaaatgc  46140
ttatttcaag aaacatgaaa gataaatcaa gatattctaa ttcccaagta agaataattc  46200
cagaagcaga aaatagaata gaggcaagga aacactcaaa acttctccag tgccatagaa  46260
atgtgtatta atctttagaa tgaaacggac taccaaatgc tgagcaggaa gaacaaaaga  46320
gatccactct taagccagtg tggtgcccaa gcgcagtggc tcatgcctgt aatcccagca  46380
ctttgggagg ccgaggcagg tggatcacct gaggtcagga gtttgagatc agtcaggcca  46440
acatggtgaa acctgtctg tactaaaaat acaaacatta gctgggtatg gtggtgcaca  46500
tctgtaatcc caactacttg ggaggctaag gcaggagaat cacttgaaac caaggaggtg  46560
aggttgtagt gagccgagat catgccacac tcccagcctg ggtgacagag caagattcca  46620
tctcaaaaaa aaaatccact cctagacaaa taatagttaa attttagaac accaaggaga  46680
aagaaaaaaa attgtaaagc ttcagagaaa ataaacatta actacaaaga aacgagagtc  46740
agacgcgtgc acttcttcct agataccagc agataaagca atatctccaa aattcagaag  46800
gttttaacgt agaatcctat acccagtcaa gaatattcac atggaaaagt gaaataaaaa  46860
```

66

```
acattgttta aacatgcaag ggttcagaaa gtttaccatt cacagaatcc ctgaaaacaa   46920
aaccaaataa tcacttaagg actcattaag aaaacaaatg aaataaaagc accaatgatg   46980
agtaaaataat cagaaaaatt tacagtttac ctaaataact gtttatgcat aatgtatgaa   47040
aacccaaaaa tttaatatgg gacagaatta aaatcatgat aagattcttt tttgctttac   47100
tcatggagag ttcacataaa cagattatct tttaatagca agagaaaaaa atgtttagat   47160
atgtgtgaaa aactaagggt accaaaacag tgcaaatcca tttatcatca ggaaaatcca   47220
aattaaaacc acagtatcca ccagaataac taaaaggtaa aagacagaaa ttaccaagag   47280
ttggcaagaa tgtggagcaa ccacatatac ttctggggta aataagttgg tgcaaccggt   47340
actgaaaact gtttgctagt atctactaaa accgagcaca tgcacagact acaaccaagc   47400
agttccactc ccagatacac actcaacaga aatgcacaca ctcactcaac aaagacagtg   47460
tactagagtg ttcatgtact tactattcat aatagtccaa aaatgcaaac aaccaactgc   47520
caatcaaagt caaatgtata tctatattag ggatatatac aatggcatat acacagcaat   47580
gagaatgaaa tgaaccagct cggcacagtg gttcatgcct gtaatctcag cactttgggc   47640
gggtaaggca ggcagatcac ttgaggtcag aaatttgaga ctagcctggc caacacggtt   47700
aaaacctgtc cccactaaaa acacaaaaat tagccgggca tagtggttgc aggcctgtaa   47760
ttccagctac tcgggaggct gggttgggag aatcgtttga acccgaaagc cggaggtcgc   47820
agtgagcgga gatcgtgcca ctgcactcca gcctggacga tagagcaaga ctccgtctca   47880
aaaaaggaaa tcaaaaatat aaaataagat gacaggaaata atccgcaaaa gatcagtaat   47940
caaaataaat ataaatgggc taaagctacc tattaaaaga caaagatttc acacccataa   48000
ggatagctac tatcaaaaa agagagagaa taaccagatgt tagcaaggat gtatggaaac   48060
tgaaattctc acgcattggt ggtgagaata taaaatggtt cagccctcgc gggaaaacact   48120
atgctgggtc atcaaaaaat taaaaataga agtactactt gatccaacaa tctctacttct   48180
gggtaatac ccaaataact gaaagcaggg tcttgaagag atatttgtac acccatgatc   48240
atggcagcat tattcataat agctatgatg tggaaccaac ataaatatcc tttgataaat   48300
atatggataa gcaaaatgtg gtgtatacat tcaatggaat attaattagc aataaaaatg   48360
aagaaaattc tgacacatgc tacaacatgg atgaaccttg agggcattac attaaatgaa   48420
ataagccagt tataaaaaga caaatactat atggggtact atattagata ctcatgcaag   48480
gtacctaaaa taggcaaatt catagagaca aaaagcagaa tggtggttgc cagggggctgc   48540
ggtaatggat acagagctt caatttttgtaa gatgaaaaaa ttctggagat tggttgcata   48600
acaatgtgca cacacttaac actggggaac tgtaaactta aaagtagtaa atggtaaaaa   48660
taaaaataat aaataataaa tttttatgtta ttttaccaca atatttatta aaagacaaag   48720
attaactaat taaacaaaat ccagccataa gctaatggta agagtaacaa ttaaagaaga   48780
cacagaaaat tgaaaatcag tgactagaaa aagatattcc atataaatgc taacaaaaag   48840
caagtacagc aatataaaga gaatgaacaa aaaaaaaatt aaataagatg gctcgtttat   48900
tcccaaaagg tacaattcac caagaaagta caagaattgt gaaccttaa gcacataaaa   48960
cagcttcaaa aatacaacat ttaaagaaaa atatatatta aacatagaaa tagtacaaaa   49020
acccctacaa gaatcataat gggagtcttc aatacaactc tccatatcaa caggtcaaac   49080
agagaaaaaa aataagttaa ggatgcagaa aacctgaatt accatccaata aacttgagat   49140
taatatgaa ctgtatacccc aatatactaa gagttcaggg aacagtcgtg actgacagtg   49200
gactgcaaat taatctgttc ttaatctttg tttttcttttc agcactgtgg cagaatagag   49260
atcctaaaaa ccttccagct acaaaacatc tttttaaaaa tataaaaaa tacaaaaata   49320
actctgaaat caatagaaga cacatgttga aaccaaaatt ctagaatcaca gggagaataa   49380
aggcattttc agatattcaca aaaacagaaa attgatcatt gctgaagtaa ttctaaagaa   49440
atgtacttga gggagaagaa aaatgttcca aagaaaagta tctgtgatac aagaaaggaat   49500
ggaaagtgaa gaaatggtaa acaggtagat aaagctaata aatgttgacc tagaaaataa   49560
caaaaacaat agcaataatg tctcgttggg agggttgaag taaaaataca attaaggccca   49620
aatgtgaggt aagtggaatg aaagaattag aagtccttgc cttgttcaca ggactgatta   49680
aataaatgag ccaggttttc cattcaaaca gttaaaactt gaacaaaata aactcaaatt   49740
aagtagaaag ataaaaaaca gaaattaatg tcatagaaaa ataaaaaatc aatagaatta   49800
atcaataaat cctggttaat aaaagctggt tctttgaaag gattaataaa ataatcatta   49860
agcaagtctg atcaaaaaaa aagagaaaag gtaccaaaaa aagtactgta tcagaaagag   49920
aacatacaga tacatacaga tatgtaagag tctgttttct tacaccagaa tactatatac   49980
aacattatgc tagcatatat taaatttcaa taatgttaat gattttctag gaaaacagaa   50040
aatattaaat ttactttgaa gaaacagaaa aactgagaaa aataaatgat catgaaaaaa   50100
atgaaaaggt aattaaaatac tgatattaac tgcctaaaca acaccagcag cagcccaggc   50160
agtctgcagt caagttctgc caaacttgag ggaacagata attcttctat tccagagcat   50220
agaaaatgat ggaaagtttc ccaatttaat cagagaggac agcctgatcc ttgttatgaa   50280
cacagataaa aatggggtaa actatatgcc aaactcagat accaaaaccc taaataagat   50340
gctagcttat tgatgtgaac aatccaaaag tgcattttaa attagcccag ggttttagag   50400
aaagaaaat tagcaatgtg accaccactt atgttaacaa ttttaagacg aaaaattaca   50460
tgatcatatc aatgcatgct acacaaaagc atttgggcaa aaaacccaac acccacccct   50520
gacttttaa actcttagta attaggcata aacagaaatg tacttaatgt gatagaataac   50580
actcggtgaa gatacagagg gaatgctccc taaaaccaag cccaagacaa agattcctat   50640
ttaacctcaa tagtcaacac tgcagcgaga gtaatctatg gaagacaagg aaaaaagtaa   50700
aaacatgaga gacatctgtt gtttaacaga caataagatc acctacttgg aagaggcaaa   50760
cgaatcaagc gaaaaactat taaaactgag acaggcttta gtatggaggc tcagcttcag   50820
ctgtagtttg ggctaccaaa ttcaactcgc ttgcttggag agttaatcct gaaaagctaa   50880
```

67

```
tttctgttga ggtattagga ttgacaagcc tgtgctcctc cctcctcccc catcttcaac    50940
actgaaataa cacggtgttt ggaactggat aacagaatct tccaaaaaca aaaattgtcc    51000
tgaagggctg acttgtgccc ttactcaaaa aacactttat ctgctgcctg cagctcctac    51060
agttgctggt ggataagcct gccaaccagc tcggcgtaat tcttcctgca gagggcaagg    51120
aagagcactt tcacaggaaa attttttttcc gaactgtatg ccgcttatta cataaactta    51180
cgtgctggca aatggagctc cagcaaaata agatattcag agtcaaactt ccttaggaaa    51240
aaaaaaaaaa aaaagcaagc acataacact aatttccttg catgggcact ggggaaggag    51300
gtcgttactt ccgcacgccc gcaggtccgc accaccggga aaccacggg cacgcgcgcg    51360
tgccccccggg ccttccaggt gcactgcgcc gcggcgcccc agctgacccg ggatgcgcag    51420
ccctagccct tccccctgtca ccccggccag gaagggcgg gagcgcggcg gagcgcgagg    51480
gcgaagggct tctcggtcct ctgcaccacg cagcacccc aaggcacaac agggagggtg    51540
cgggaggctc ccgagaccca ggagccgggg ccgggcgtgc ccgcgcacct gtcccactgc    51600
ggcgagggct ggggtcgcct ccagggccgc agctgtcggg agccacctgg ctctcagtcc    51660
cgggtccctg cgacaaccct cgggcccgga gggagggagg cggccaccctg ccgctgccac    51720
ctgcggcacc ggtcccaccg ctccgggccg ggcaggacag gccaggacgt ccctcctggg    51780
ctggggacag gacacgcgac gagggaccg gggccccgc ggcgaagacg cagacgcagg    51840
tcccagaaag gcagtcccgt gccccccacga cggactgccg gaccccgcg ctcgcccgcc    51900
catccttca gaccacgcgg ctgaggcgca aagagccggc cggcgggcgg gatgcgagcgg    51960
cggctagtac tcaccggccc cgctggctca gcgccgccgc aacccccagc ggccacggct    52020
ccgggcgctc actgatgctc aggagaggga cccgcgctcc gccggcgcct ccagccatcg    52080
ccgccagggg gcgagccgca gccgcgcggg gctcgctagg agatgtagta ccgggaccgc    52140
cgcctgcgcc gtcctcctc agccggcggc cggggggccc ctctctccca gtctcagtg    52200
tctcatctcc ctatctgctc atcctctcgt cgcacataat cgatgtttgg gcgtcccaag    52260
ccagatgtgg accccatttc cgcactctac actggaggtt ttctaagggt ggtgcccgga    52320
ccagcagctt cagcctcatc tgggaacttg agaaaatgca gattctccgt cccacccagc    52380
ctattcggtt tttcctgcac taaaaccatg aaggtggggc ccagcagtcc acattctcgc    52440
aagccgtca agtgattctg aggcgccctc cagtttgaga gctatgctca cggcctcacc    52500
tccgccccgc aaggagcccg gtcttgcctg tggcgctagc cgcacacgga cacctcatcc    52560
tgcggggccc gcccccccgc tgcaccctca ccgcccaacg cctcctccgg gatgcagcgg    52620
aggcgcctgg aagtcggcaa ggtcaacatc ccctcagca tcttccctac cctcacggct    52680
cctcctccag gggtgcctca tggccaggg ttagaaagag ccactgtgtt tcttgacatg    52740
gaagtggcct aagaccttaa tgaaaactgc aggagtggaa tgacagaacc tttggtcata    52800
cttgagggcg tgaagctcaa atggaggaga aggaaagat ccagggagaa taaccaaccc    52860
tggcaagttg tggcgcccag gtagaggggc gagcctaggc tagcggttct cgaccagggc    52920
cggtgttgcc cctcctcgcc gcccgcgta catttgggga ggtctggaga cattttttggt    52980
tgtcatgatg cgggagttgc tactgttgcc taagtgggta gacacgaggg tgctcctcaa    53040
catcctacct gaaggacagg actgccccac aaggaagaat gatccggccc caaataagaa    53100
accctgggct ggtcagcaac aaccccttttg ttctgagaag agaggaggaa agaataaaag    53160
aagtggggtg aagttttggt ttggtagagg aaacttgaag acattttcac tggaagagcc    53220
gagaggaaga ggaggagagt gtctgtaagg acgagcaaac cgggtgacag ctgatttcct    53280
catattgaag taatgagtcc tagttataat aaattcctaa taaaaaccca gtttatccct    53340
gcaataaact tgtctttttt ttttaaatat actgcttgat tctgtttgct aatattttat    53400
ttacaggcctt tgcattgata tgcaaaaatg agatgggcaa taattttctt tttgaatgtc    53460
taatgtttgtt tggtttcaga atcaatgtta tgctcacatc ataaaaaatt tggaaccgag    53520
gcaggaggag tgcttgaggc cagaagttcg agaccagtct aggaaacaca gtgagaccc    53580
cccatctcta caaaaaaaaa aaaagaaaaa aaaatgggca tgtttgcttt ttcctttcac    53640
tctgaacaat ttaaggagca ttaaaattat ctattcttttg aggtttgatc atttcccagt    53700
taaaaatgtt cctcccagcc tgatgcttttc tttggggagg gtaaatcttt taaggctaga    53760
aaagtttctt ctgtggcaat tttattattt acattttaaa aattattcta gagttaattt    53820
tgataaagca tgtatttctt aaaacaaatt atccttttttt tccagatgct caagtgtatt    53880
tgcataaagt tgaggaaagt agtctttttgt gaatcttttca acttctccca aatatcttat    53940
tttgtgtaatt tttgcttctt tatttttgtta actttaaaa gtgtattttt ttttcaaaga    54000
atcagctctt aggtttatgt ttttggttat actggagctt ttttcttctt ctttttaaaa    54060
tatttttttct ccttttatttt ttagacgtat tttgatctaa cgtaatcgga agaaggtaaa    54120
ttagaatctt ttgttactat tgtgttttta tttctccttta tttctctgaa gtcctgcttt    54180
ataaatagta ccatgttatt tgtgcataaa tattcatttg tcttatattc ttttggacttt    54240
tcccacttca tcataaaatg accttccttg tctcatttaa tgtgttcaaa ctttgccctg    54300
aatttaactt tgtctgatat tttaccatcc tgctgaattt tgtttgttac cccaacaccat    54360
ctttgctgtt ttcgtctttt ctgaaccctt tattttaggt aatcccttga attagagcac    54420
taagttttgc tttgtgatta aatctgaaaa tctttatctt gccatagatg agttgagccc    54480
tattcatgtg acagctatat tatgctgttt catagccctt ttggtccttt tttcactctt    54540
gcattgcata ttttgtgttt attgtgtttt gtgttttctc tgataatttg gaaggtttgt    54600
atttttattc aggggagttgc cttataatca tactccgcaa tacacatcgt cctcagtttc    54660
ttcagactgt ctgttaactc cctattctga ataaaaatga cattgtaatt tccctctttt    54720
ttctttaccc cttttcttct cctcacctaa tgtaaatgat tttatccttc tttagtatttt    54780
gctttttttaa ttaactacat ttataaatat ctttatcact tgattttttaa atcagctttg    54840
aatgagatat ttggattcct agatataaaa gatgttaatt ataccatttc cacgttagta    54900
```

68

```
ggtttataaa atcatacatt ctgctgtgta accataatcc cacgtttgtt ttagttccac   54960
tcctacagtt aaaagattca gaagtattat taacagttat tttgccatag tttttccccc   55020
aacccatttt gtggtaagtt atgatcctgc tttagtttct taagaataat ttatagagca   55080
gagtgtggtg gctcacgttt gtaatcccag cactttggga gacaagaggt agaaggatcg   55140
cttgaagcca gcagttcaag accaccctga gcaacatagt gagaccttgt ctctacaaaa   55200
aattttaaaa tttagccaga cgtagtggcg tgtgcctata gtcccagcta ctcaggaggc   55260
tgaggcaaga ggattgctag agcccagaag tttgaggctg cagtgacctc tgattgtgcc   55320
actgcacccc agtctgggca agaaagtgag aacctatctc tttaaaataa caataataac   55380
ttatgaaaat tatattccct gagtttttca tgtttaaaaa tatttgttgc ctttatcctg   55440
taaaagtttg agtataaatt cttgggttat actttatttta ttgaagaatg taacactttg   55500
gtcttctaga attgagtgtt gctgtaatga aaccagaagt cagcctggtt tattttttcct   55560
cagaaatgag gtaattgccg gccggacacc gtggctcatg cctgtaatcc caacactttg   55620
ggaggccgag acaggtggat cacgaggtca ggagattgag accatcctgg ctaacatggt   55680
gaaaccccgg ctctactaaa agtacaaaaa gttagctggg catggtggtg gacgcctgta   55740
atcccagcta cccgggaggc tgaggcagga gaatggcgtg aacctgggag gaggagcttg   55800
cagagagctg agatcgcgcc actgcactcc agcctgggcg acagagtgag actccgtctc   55860
aaaaaaacaa aaaaaaaaca aagaagtgaa gtaattgcca tgatgctcca agaattatct   55920
ctttgtctat gaaatccaag aatctcactg ttatacattt tggaattatt attctgggcc   55980
aatattttcct gggacacaat agattgactc tataagattta attttttttt tttttttgag   56040
acagagtctc actgcaatct cagcttactg caacctctgc ctcacgggtt caagtgaattc   56100
tcctgcctca gcctcccaag tagctggagc tacaggtgcg tgtgcaccatg cctggctaat   56160
tttttgtcttt ttagtagaga caggtttca ccatgttggc caggctggtc ttgaacgcct   56220
aacctcaagt gatccacctg cctcagcctc ccaaagtgct gggattacag gcgtgagcca   56280
ccatgcccag cctcaattcc tctttctatc tggtaatttt tctgaagttg aaaacatttg   56340
ttctaatacg ttatttcagt gttcttctaa gatgtgtaaa gacccctatt cccaggtaaa   56400
ccccatcctt gctagtgagc tcggctggtt cttcacaaga gctctggttt tctcctgctt   56460
aatctcaagt accctgtca gcctccacct ggtttatgat ttggagtttt ttggtttttg   56520
ttttttgttt ttgacagagt cttactctgt cacccaggct ggagagcagt ggcataatct   56580
cagctcactg caacctctgt ctcccaggtt tgagcgattc tcctgcctca gcctactgag   56640
tagctgggat tacaggcgcg tgccaccaca cccggctaat ttttgtattt ttagtagaga   56700
tggggtttca ccatgttggc caggctggtc ttgaactcct gacctcaggt aatccacctg   56760
cctcagcctc ccaaagtgct gagattacag gcgtgagcca ccgcgcctgg catggtttgg   56820
agtttttaatc tgtagttttta ataaagatag tgcttatgtt tgtgtttctt atatttcttg   56880
gtactcttgg gtaatttgta agatcccat atctacacaa gaagtccatt ttcaattctt   56940
ttcttcagac tgtttatttt attttattt atttttattt tatgtttgag atgggtgtct   57000
gctgtgtcac ttctggaggc tggagtgcag tggcgcgatc tcaggtcact gcaacctccg   57060
tctcccgggt tcaagcaatt ctcctgcctc agcctcccga gtagctggga ttacaggcac   57120
ctgccacttt ttaattttt tagagacaga gtctcgcttt gttgaccagg ctggagtgcg   57180
gtggtgcaat catggctgac tataacctcc aaatcctggg ctcaagtgat cctcctgcct   57240
cagcctcctg agtagctggg actacaggca catgccacca tgcccagtta atttaattt   57300
ttttgtagag acagggtctc catatgttgc ccaggctggc ctcctactcc tggcctcaag   57360
taatcctcct acctcagcct cccaaattac taggattata agcatgagcc accatgccca   57420
gccttgttct actactttaa tttcatatgt taggtgacca tgtaattgat catccaaacc   57480
aggatactgt aagaataaaa gaggctgaca gtagtatgat gctgggacta gcattgtgca   57540
ctgagattat ttctgggaaa gcaggagata cggtcaccct acttatatgt tgcttgtctt   57600
tggattgtag aatttggagt ttctattgtc aggcttattt caactgggca gccttgatcc   57660
gccctgccca gcaatgctac cgttctctcc accgggtctc tgggaccct tcagtcacta   57720
tacttagctc agttcccac cctcccactc cctaaaagcg taaccaggaa tcctgcctca   57780
ggtctactgc cgtcttccgt gggctgtttc agttcctatt acccagagtc aaactcccag   57840
cattccctac ctgattccag acttggagtc cagagcttta acctcttcag gccaactccc   57900
cactttgcat ttctgtccct atatcttagt ccatggagat acatttcatg tctttgagtc   57960
tacttacaaa gtaaattttg ctgttttta attttttttt tgagatggag tcttgcctg   58020
tcacccaggc tgtggtgcaa tgacgccatc tcggctcact gcaacctccg cctcctgggt   58080
tcaagcgatt catctgcctc agcctcccaa gtagctgtga ttacagacag gcgccaccac   58140
gcccagctaa tttttttttat cttttagttag agacagggtt tcaccatgtt gacgcaggctg   58200
gtcttgaatt cctgacctcg tgatctgccc atctcggcct cccaaagtgc tgggattaca   58260
ggcgtgagcc actgcgccca gccaattttg ctttttttat atttcattgc tatatgttta   58320
gaggataagt ttacagtgct atatgcattc ccaaatatta gaccaaaaaa atctccaaaa   58380
aattagaaag aaaatccaaa aaatctcaaa aaataccaaa aagcaacaat ctcacagacc   58440
atactcactg acccccaata aaataaaatt agaaattaac cacaacttaa caaaccttaa   58500
tactcaagtc agagaggaaa gaggaaataa acatcaaaat tacaaagtct aggcggtggc   58560
tcacgcctgt aatcccagca ctttgggagg ccaaggcggg cagatcacaa ggtcaggagt   58620
tcgagaccag cctggccaat atggtgaaac cccgtttcca ctaaaaatac aaaaattagc   58680
caggcatagt gatgtgtgcc tgtaatccag ccacttggga ggctgaggca ggagattgct   58740
tgaacccagg gagacgaaga ttgcagtgag ccaaaatgct gccactgcac ttcggcctgg   58800
gtgacaaagc gagactccat ctcaaaaaaa aaaaaaattac aaactctttta gatagaaatt   58860
ttggtgtttt tttttgagac ggagtctcac tctgtcgcag aggctggagt gcagtgggac   58920
```

                                   69

```
tatgtcagct caccgcaacc tccatctcct ggattcaagc aattctcctg tctcagcctc   58980
ccaagtagct aggattacag gcgcccacca ccagacccag ctagtttta tattttttagt   59040
agagatggtg tttcaccatg ttggccaggc tggtctcaaa ctcctgacct caagtgatcc   59100
acctgcttca gcctcccaaa gtgctcagat tacaggcgtg agccaccgca ccccacctag   59160
atagaaattt caacatgagg ccgggcacaa tggctcacgc ctgtaatctc agcacttcag   59220
gaggctgagg cgtgggagga tcacttgggc ccaggagttc aggaccacga tggtgaccag   59280
agacagaccc tgtctctatt tatttgaaaa aaaaaaaaaa aaagagagag agaaagaaat   59340
ttcaacatga aaagtatctc tcaaaccctt cgagatgttg gcaaaaagcg actcaaagga   59400
aaatgtatta ctgtgtgtga atttgcttga aaataagaaa gaggccgggt gtggtggcta   59460
acacctgtaa tcccaacact ctgggagtcc gaatcaagtg gatcatgagg tcaggagatc   59520
gagaccatcc tggctaacat ggtgaaaccc tgtctctact aaaaatacaa aaaattagct   59580
aggcgcggtg gctcatgcct gtaatcccag cactttggga ggctgaggca ggtgatgca   59640
ctgaggtcag gggtttgaga ccagcctggc ctacatggtg aaacctcgtc tcttctacaa   59700
atacaaaaat tagctgggcg tggtggtggg tgcctgtaat cccagctact caggtgctga   59760
ggcaggagaa tcgcttgaac ccgggaggcg gaggttgcg tgagccgaga tcgcaccact   59820
acactccagc ctgggcaaca gcctgggtga cacagtgaga ctccatctca aaaataacaa   59880
aaaattagct gggtgtggtg gcctgcgcct gtagtcccag ctacccggga ggctgaggca   59940
ggagaatgga gtgaacctgg gaggaggagc ttgcagtgag ccgagatccc accactgcac   60000
tccagcctgg gcgacagagc aagactcttg tctcaaaaaa aagaaaaaa aaggaaaaaa   60060
gaaccctgat aataaagaaa ccaaatgttc aactctcaaa gctcggacac ttaaagaaa   60120
taattaataa aggcagaagt taaagggagg atgataaagc aatttttttt gttggttttt   60180
ttgagatgga gtcttgctct gtcacccagg ctggagtgca gtgatgcgat cttggctcac   60240
tgcaacctct gcctcccggg ttcaagcaat tctcctgcct cagcctcctg agtagctggt   60300
actacagggg cgcgccacct ggccagcta atttttgtat ttttattaga gacggggttt   60360
caccatattc gttaggctgg tctcaaactc ctgacctcag gtaatctgcc cacctcggcc   60420
tctcaaagtg ctgggattac aggcaggcgc caccgcgcct ggcctaaagc aaaatattgg   60480
ttctgtgcaa aagtcaata aaaagagcaa acgtttacaa actggagcca gcacccattc   60540
agctcagtgt gtctgagaa aaaacaatct cgcttcagaa ttcatgatta cgcagcccttt   60600
tttgcttcct aaaaatccta ctatgttgct gttgaccatt ctctctcttt ctctctctct   60660
tgctttctct ccagaaaagc tattcagaca ttctcctctt tcctcaaacc tccaacactt   60720
cctctccat ccttagcctc agctgctgac ctcacttcta atcattgaa aaccaggaga   60780
agcatttaag agtgaacctc cgcctccccg cacgggcaaa accacccacc cacagaattg   60840
tgccccaatt ctgcgtcctc tcctctcacc atggatggac ggtccaggct ccgagccaaa   60900
gccaggcctc ccctggagct ctggatccac cacctgcagc ttctcaggca gggcccagc   60960
agctcccctg ctcccttgta ccatcaatcc ctcccctcac tgggtcactc ccaacaatat   61020
atatatttag tgatgtttct ccatgtggt aaaatcactt agcctctctc ctcccccagc   61080
tactatccta tttgtttctt tccattctct gcaaaacttc tcaaagcatt gtgtctatgt   61140
gctgactcca tttatcttct cccgttctct gctgagtcct tcccacagac tctcacccca   61200
gttactccat gaaatgacct ctgcactgcc acatccaatg gtgaatgttc agttcttaat   61260
tttattcagt cttttcagcag catttgacct ggccgatcac tccctcttct taaaaatact   61320
tttctcagcc aggcgtgatg gctcacacct gtaatcccaa cactttggga ggccaaggcg   61380
ggaggatcat gagagcccag gagttcaaga tcagcctggc caacatggca agaccctatc   61440
tctacaaaaa ctaaaaagta gccagtgtga tggcatgcac ctgtagtccc atctacttag   61500
gaggctgagg cagtaggata acttgagccc gggaagtgca ggctgcagtg gactgtagtc   61560
gcaccactgc actccagcct gagtgacaga gagacctgt ctcaaaaaga caaataggga   61620
aactttctct agcatatcc tctgattctc ctgctgcttc tgtctgcaca gattcagtct   61680
ccttgccgg tcttctctca tcctcctgat ctcttgacct tgaagtgcc cagagtacag   61740
tcttttttt tttttttgag acgcagtctc gtctgtcacc caagctggag tgcaatgggcg   61800
aggtctcagc tcatgcaacc tctgcctcc gggttcaagc gattcttcctg cctcacggcg   61860
ccaagtagcc aggactacag gcacatgcca ccatgcccag caaattgttg tatttttagt   61920
agagacaggg ttttactata ttggccacgc tggtctcaaa ctcctgaact cgtgaaccac   61980
ccgcctcggc ctccaaagt gctgagatta caggcatgag ccaccaccc cgcccagag   62040
tacagtcttt agacggcctc tctacctata cttgctcccc tcataaactc ctcctgcctc   62100
atggctttaa ataccatcgg tagactgatg actcccatat ttctctttt tttttggaga   62160
cggagtctcg ctcagtcccc caggctggag ttcagtggcg cgatctcggc tcactgcaag   62220
ctccacctga caagttcaca ccatctctcc acctcagcct ctccagtagc tgggactacag   62280
ggcacccgcc accacgcctg gctaattttt ttgtattttt agtagagatg gggtttcacc   62340
atgttagcca ggatggtctc gatctcctga cctcgtgatc tgcccgcctc ggcctcccaa   62400
agtgctggga ttataggtgt gagccacccg t gcccagccga tgactcccat atttctatct   62460
cttgctgtgt gggagttctc ctcagaactc catctcata aatccaactc tcataaatag   62520
tatctcaaat gggcaatatg ctcaaagtc aattcctact tttctcccta aacttgcttt   62580
cctgcagtct ccaccatctt aatgtccaat ctaacattag gaggcaaaaa ctttgaagtc   62640
attcttgactt cttctctatt acacaccta tccaatcttt ctgcagatcc agtcgaccc   62700
caaatccagt tagctctcat catctccct gttaccccct ggtccaggcc atcttcctct   62760
ctcacctgaa tcactgcagc attctcctca ctggtctctt tggttctgtt ttcactccac   62820
cttagcatag tctccacaga gcagtcagag ggatcctttt aaagtgtaat tcccatcctg   62880
tccctgctct gctcaaaacc ctgtcgtgat tcccgtttta atctgtcaga ttaaaagcca   62940
```

                                    70

```
gagtctttcc agtgacctac atgatctgcc tattatcacc tcccacttct ttccccttgc    63000
tcactccact ccagctctgc agctgtcctt tctgtttcct gaacagccca gattttgctt    63060
ctttagaacc tttgtatttg ctgtcccctc tgtctggaat gttttccag gaagtcacct     63120
ggctctctcc tgcacttcct tcctgaccac catgtttaaa aatcactcaa acacacttca    63180
ggccggacat ggtggctcac gcctgtaatc ccagcacttt gggaggccaa ggtgggtgga    63240
tcacctgagg tcaggagttc gagaccagcc tggccaacat ggtgaaactt cgtctctact    63300
acaaatacaa atagtagcca ggtgtagtgg cacacacctg taatctcagc tactcaggag    63360
gctgaggcag gagaatcgct tgaacccaga aggcagagga ggtgcagtga gccgagatca    63420
cgccacaaca cccagcctg ggtgacagag caagacccca tctcaaaaaa aaaaaaagaa     63480
aaaaaaatca cacaaacaca cttctcttca tattccttt ccaagtttta ttttctctcca    63540
gaatacttta cattgtttta atggaagttc tccgtttccc cccaactaga atgggatactt   63600
cctgcaggta ggcactctag tcctcccatc caagtactaa ccaggctcaa ccctgcttag    63660
cttctgagag caggggagat caggcctgtt caggtggta tggcccagga attttgattc     63720
tgttttatca attgctgttc tgttgattcc cttttgttcc tcctcctagt gctgagaaca    63780
ctacttgtac ataataagca ttcaataaat atttgttgga tgaatgactt gttgaatgaa    63840
ttaatctgaa aaatgcagga ctggttctac attagaaaat ttttcaaggt cattctctgt    63900
tgtcgtaaca cattaagaga ggaaaatttt gtactctaaa tcatttgata aaatacatac    63960
tgatttctgt tttcaaaaac tcttagtggc tgggcgaggt ggctcacatc tataatccca    64020
gcattttggg aggacgaggt gggcggatca cttgaggtca ggagtttgag accagcctgg    64080
ccatcatggt gaaaccctat ctctactgaa aataagaaaa ttagccgggt gtggtggcgc    64140
cggaggttgc agtggctcaa ctgggaggct gaggcaggag aatggtcca acccgggaggt    64200
ctccatctca aaagaaaact cttagtgagt ttaggaatcc aaggaagacc ctcaaactca    64260
atagataatc tagctaccag aagccttcag taaacctaa cactccatgg tgaaacatta     64320
gaaacattcc tactaaaaga caggctaaga atgcctgcaa tcttcacggc tagtccaaga    64380
agtcaaaaag aagaaatgag cgctgattta aaaaaataaa caaacaaaaa actaccgatg    64440
cagaggctgg cagcaaggac tgaaggactg tacagtactt gcctggagca ggcggatggc    64500
cacacccctg cgaagcctgc tcagctggct gggggacgct ccagtgtgtg agtggcagga    64560
tgcagggtac ttcctctgcc agggagttgc actggggaga tcctccccca ctcacacttt    64620
ggcagctggg gctttggaat gtgacttagc ttctgtcaaa gggtcaatcc acctttgat     64680
atatgatgca aaggcgaaca tatgatgcaa aggtgagaga acagcccaaa ttaggacttt    64740
taccacagct gtggaggtgg acagcgacag tggtgggccc tggccagact tttcatgtctc   64800
aaaggtggtg gttgttctc ctacttcttg tccctccagg gcttcctttg cctgtgtgct     64860
gaacctgctt ctttaattt tttttaactt tttaaattt ttaattgttt taattaaaac     64920
aaattttgaa aactgtctga acctgcttt gaaccctgct atgatttgaa tgtttgtccc     64980
ctgccaaact gattttgaaa cttaatctcc aaagtggcaa tattgagatg gggctttaag    65040
cagtgactga atcatgagag ctctgaccc atgagtggat taatggatta atgagtttgtc    65100
atgggagtgg catcagtggc tttataagag gaagaattaa gacctgagct agcatggtcg    65160
cccttcacc atttgatatc ttacactgcc taggggtcct gcagagagtc cccaccaaca     65220
agaaggctct caccagatac agctcctcaa ccttgtactt ctcagcctct gtaactgtaa    65280
gaaataaagt cctttcttt atgaattacc cagtttcaga tattctgtta taaacaatag     65340
aaaacgaact aaggcaaact ctcatgattc tactgccatg ccattccaat aaactcctt     65400
tatgcttaag agagccagag ttggccaggc gtggtgactc acgcctgtaa ttccagcact    65460
ttgggaggcc gaggcaggtg gatcacaagg tcaggagatc gagaccatcc tggctaacac    65520
ggtgaaacct gtctactact aaaaatacaa aaaattagc tgggcgtggt ggtgggtgcc     65580
tgtagtccca gctactcggg aggctgaagc aggagagaaa tggcgtgaac ccaggaggcg    65640
gagcttgcag tgagtcgaga tcgtgccact gcactccagc ctggtgaca gaagtgagact    65700
ccgtctcaaa aaaaaagaga gccagagtt atttctgttg cttgcaaca agaaatctgg     65760
ctggtgcact gaagtttcca taaataatag caatttaaag actttttcca agccaggcaa    65820
tgcctagcct tgtgtgagtcc ttgtggtaat acattcattc attcatttgt tcaaccaact   65880
gtgctccgaa gactaagaat acaaaaatgg gggccgggtg tggtggctca cacctataat    65940
cctagcactt tgggaggccg aggcaggtag atcacctgag gtcaggagtt cgagaccaac    66000
ctggccaaaa tggtgaaacc cctactctac taaaaataca aaaaattagc tggggcgtggt   66060
ggcggcacacc tgtaatccca gctactcggg aggctgaagc aggagaatcg cttgaacccg    66120
ggaggcagga gttgcagtga gccgagatcg caccactgca ctccagcctg ggcgacaaga    66180
gcgaaactcc acctcgaaaa aaaaaaaaaa aaaaaaagag ggccgggcgt gggcgcaggtg   66240
gctcactcct gtaatcccag cactctggga ggccaaggca ggagaattac gaggtcagca    66300
gatcgagacc agcctgacca acatggtgaa accccgtctc tactaaaaat acaaaaatta    66360
tccgggcgtg gtggcgcaca cctgtaatcc cagctacttg gggaggctgag gcaggagaat    66420
cgcttgaacc cgggaggcgg aggttgcagt gagccgagat catgccactg cactccactcc   66480
attgtagagt tatagtgtgc ttctagtttta gttgagagaca atcaatccag cagtcaaaac   66540
ctagaatcta taccctgagt ccttactgaa atcaatccag cagtcaaaac atgggaccaa    66600
cgatcacgac agtagagatag gaagagcacc tttgtacatt tagtcatgt tgagatcagt     66660
cactgacaga gctgaaggaa gctcacagtt ctggggttca tcctttggca tttaaaaaga    66720
aaagtgctaa gaaaattcgg ttggtcacgg tggctcacgc ctgtaatccc aacactttga    66840
gaggccaagg caggcagatc acgaggtcag gagttcgaaa ccagcctggc caacatggtg    66900
```
                                    71   66960

```
aaacccgtc tctactaaaa acagaaaaat tagccgggca tggtggcgca tgcctataat    67020
cccagctact caggaggctg aggcaggaga attgcttgaa cccgggaggg ggaggttgca    67080
gcgagtgaga gcaggccact gcactccagc ctgggagaca gagcaagact ctgtctcaaa    67140
aaaaaaaaag aaaaaaagaa agaaaggaaa aaaagaaaga aaaaaaaga aaaaaaaga     67200
ttcaggccag gccaggcctg gtggctcaca cctgtaatcc caacactttg ggaggctgaa    67260
gcgagacggt gccttagccc aggagtttga gaccagcctg agcaacatag cgagaccctg    67320
tctctataaa aaaaaatttt tttttggcca gacgcagtgg ctcacgcctg taatcccagc    67380
actttgggag gccgaggcag gtggatcacg aggtcaggag atggagacca tcctggctaa    67440
cacggtgaaa cccatctct actaaaaaat acaaaaaatt aaccgggcgt ggtggcgggc    67500
gcctgtagtc ccagctactc gggaggctga ggcaggagaa tggcgtgaac ccgggaggcg    67560
gagcttgcag tgagccgaga ttgcgccact gcactccaga ctgggagaga gtgagactcc    67620
gtctcaaaaa aaaaaaaaaa aaaaaaaaat taattgtcag gtgtgctggc atgcagctgt    67680
agtcctagct actcgggagg ctgaggtaag aagatcgctt gagcccagga gttcaaggct    67740
gcagtaatag tgcctctcac tctaccctgg gtgacaatga gaccctctct caaaaagaaa    67800
gaaaaaaggg aaagaagaaa agaaagaaag aaagagaaga aaggaaggaa gaaagaagaa    67860
aaaagaaaag gaaggaagga agaagaaaaa aagaaaagag aagaaaggag agaagagaaa    67920
aaagaccaaa gggtcaggat cccaaaatag tttttatgtt ttatttatttt atttacttat    67980
ttatttttga gacagtatgg ctctgtcgcc caggctggag tgcagtgatg cgattggcggc    68040
tcactgcagc ctccaaactg ggctcaggtg gccctcccac ctcagcctcc cgagtagctg    68100
ggaccacagg cgcgtgccac catgcccagc taattttttta attctttgta gagatgaggt    68160
ctctatatgc tgcccaggct ggtctcgagc tcctgggctt aagccatcca cccgcctggg    68220
cctcccaaag tgctgggatt acagaagtga gccaccgcgc ctaatcgggt ggtttgtttg    68280
tttattgacg ggtctccgct gctgcccaag ctggagtgcc agtggctgtt cacaggtgca    68340
gtcctggagc attgcatcag ctcttgggct ctagcgatcc tccagagtag ctgcagctgg    68400
gattccaggc gcgccaccgc gcggggctca gaatgggttt ttatattgag ggttatgctg    68460
ccacctagag gatatatgta gtaccgaact gtgtgcgcag ggaggctgag gttgcagtga    68520
gccaagatga tgccagggca ctccagcgtg ggtgacagag caagatttca tctcaaaaaa    68580
aaaaaaaaaa aaaaaaaaaa aagaattgaa agtaaggtct tgaagagata tttgtgcctg    68640
tatgtcata gcagtattaa ctttgaccca ctagctaaaa cacaaaagca acatgtgtct    68700
gtcagcaggt gaacggataa acaaaatgtg gtatatatgt acaattgaat attattcagc    68760
cttttaaaag gaataaaagg ctggatgcgg gggctcacgc ctgtaatcct aacacctttgg    68820
gagactgagg tgggtggatc acccgaggtt aggagtttga gaacagcctg gccaacatgg    68880
tgaaacttca tctctactaa aaatactaaa attagccggg catggtggca cttgtctgta    68940
atccaagcta ctgggaggc taaggcagga gaattgcttg aactcaggag ccggaggttg    69000
cagtgagcta agatggccac actgcactcc agcctgggca acagagtgag actccatctc    69060
aaaacaaaca aacaaaaaat tattatttcc aaagaaacaa gaccctgggt ccatttccca    69120
gcccacacct gatgttgact cacaacacac agcctggttt gctatgagcc tgcttcattt    69180
aattgtcacc ttaacttcac atcaccctca agtcctggaa taactctttg ttgacctttg    69240
tgtgctgaag catctccatg tcgctcaacg tgcagtccct ctcactgcac tgagtcaata    69300
gccagacgtg gtctgactgc agggtcatcc ttggtggctt aggctgactc gggcatgca    69360
gggtgctctg agacctcacc gcataiaggc tttgcccccca ataaactcta taaaattta    69420
atattatgtg gtctgggtgt gtgtagcttt gcactgtctt ctcgtgacag tgccctcaac    69480
ctctttccca ggatttcctc ctctacctcc tcaagtccca ctgctctgca aagacacatc    69540
gctgcagagt ccagctccc tcctttacac cccacgacgc agcctcctct ctcagaaccc    69600
tttaaacaga gtcttttact gcagatccca agaacagcca cacccctct tccctcatctc    69660
tccagacaca cccaggtaat tatagcaaco aaggtaacta tgtagatgga gtccctggaa    69720
catgtggata gtgcccctg ggagtatgca aaagcaacat tgctggcacc tgcagagaac    69780
agggtgacat ccaggaatca gagcatgggc ctctgggagg tagggatgtg gccaggcagg    69840
ctgccaaaaa ttggtagagc aaggccacag gatctttctg accttccttc caaacagagg    69900
ctcctgtact ggtgatccct gtgttgattg accactccct tcctggggt cgtggtctct    69960
gtcccagttg cccggacttc tgtgagtgtc ctactgaggt cctttttcatg agaagcatgc    70020
tgtccttcca cctgctgagg gcaagagtga caacttcaat actataatag cagtggcata    70080
cagagaagaa gaaagatgaa gtggcaagaa aaacaggctt ccaagcagga gttttttctat    70140
aaaaacaaaa acgtttacaa gcaaacttttt tataaagggc tagatagtaa atattttagg    70200
ctttgagagc cacatagact tgtttgcagg gactcaatgt cgctattgta gtttgaaagc    70260
agccatcagg gttatgtaaa tgagtgagtc tgatttttgtt tcagcaaaat tttatttacc    70320
aaaacagaca atgagtgggc tggatttggc ccatgatcct tagtttgcca actcctgctt    70380
tgggctcacc cagatctgat tttgaattct ggctctgcta ctggttagct gcagggagctt    70440
ggaaggctct ctgagcctgt ttcctcatct gtaaaattaa agcaataatt tctaacactc    70500
aagagtgtta cctcacgcct gtaatcccag cactttgagg gctgaggcag gcggatcacc    70560
tgaggtcaga agttcaagac cagcgtggcc aacgtggcaa aaccctgtct ctactaaaaa    70620
atacaaaaag tagccgggca tggtggcgcg catctgtaat cccagctact tggacacctc    70680
aggcagggat actgctagaa cctgggaggt ggagcgtgca gtgagtggag atcacacctc    70740
cacactccag cctggccgac agagcgagac tccatctcaa aaaaaaaaaa aaaaaaggc    70800
ttagaaggtt ttgagataat gaataaaaga tgccttgtgt atactaagta ttcaacaact    70860
gatagctgca ttggtctaat tataacagtt tagaagcgat tgagtcaaca aatgctggat    70920
ttgtcaggga ggacttccta tcaggaggta gatcttgggc tgagtcctga agcaaagata    70980
```

72

```
ggcattggat agaggagttg agagaacacc ctaggactgt tattattatt attcgacacg   71040
gagtctcttg ctctgtcacc caggctggag tgcagtggcg cgatctcggc tcactgcaac   71100
ctctgcctcc caggttcaag cgattctcct gcctcctaag tagctgagac tacaggtgtg   71160
tgccaccaca cccggctaat ttttatattt ttagtagaga cagagtttca ccatgttggc   71220
catgctggtc tcgaactcct gacttcaggt gatccacccg cctcagcctc ccaaagtgct   71280
ggaataacag atgtgagcca ccgcacccag cccagaacca tttttcaatc cttggctcctg   71340
ccttttatta gctgcaagat ctcaggcaat ttatttaacc tctccaaaga ctcattttct   71400
cattcacaaa atgaggcaaa taataatatc tactatccca ggttgtcatg agaattaaat   71460
gcaacatgac atttaatgaa atgagaagtc ccttggacat taactggcta aagtatgtgc   71520
tcgacaagga tatcattta ggtggatact tagcatctca gaactgatgc tcacaatgga   71580
atatcattga aacgcattaa aattcatttt aaatgattgt aggtagtgag gcaattgaaa   71640
gaagaagaca agaggactga ttataatgct tcaggctcac tagtctcctt ttaggaggga   71700
aaaacaattt caagttaaat tttaggctct agatttttac ccctgctgct cattagaatc   71760
acccagattg atgaaatcag agcccatctg aggctgtgtt tttcatctcc agacatgatg   71820
ctgttgtggg gattaagttt ttgaaaaagt acatctaaca ggtgatcgaa aatgatagtg   71880
atattattgc agtgatggtc attattgttg ttattattat actgaaagag gcttcagtttt   71940
tctgatccat aaagtgaggg aattgcatga gaccattgct aagattcctt ctagctctgt   72000
tttttttgttt ttgttttta gacagagtct ctgtcgccca ggctggagtg caatggcatg   72060
atcttggctc actgcaacct ccgcctcccg ggttcaaatg atcctcctgt ctcagcctcc   72120
gaagtagctg ggactacagg cacacaccac catgcccagc taactttat attttaata   72180
gaggtggggt ttcaacatat tggtcaggct gtctcaaac tcctgacctc aggtgatcca   72240
cccgcctcgg cctcccaaca tgctgggatt acaggcgtga gccactgtgc ccaaccccttt   72300
ctagcttttct tgatcactga ttctagggtt ctctgctgaa atatattgga gacatcctgg   72360
ataaaagatc atgcaagagc tcccaatatg gtattaataa ttgattctgg aggcttagct   72420
actcctgatg gattagacat gactcaactg cctctcttat gtgtacaaca caacaacaca   72480
accaagaaag gttattctgg cattccattt attcagtttta tttacagccc ttacttccag   72540
cagcacgtta aagatatggc cagggccggg tgcagtggct caagtctgta atcccaggac   72600
tttgggaggc caaggtgggc ggatcacaag gtcaggagtt tgagaatctg gcaattcttc   72660
agacttagaa gcaaccagct cgataacaca gtcttgtgtg ggctctccct ctgtccctcc   72720
ctcgcttccc tcatttctca tccctgcccc tgagactgtg caccttcaca tagccctgcc   72780
atgagacctt catctcaggc tttgcttttct ggggtaactg aggctaaaca ctgagtggcc   72840
ctaaaagagg attgggatt ggaagttaga ttattcacca gagaacagac tttgctgatg   72900
atcaggccca ggttgtaatt gttgaaaaaa agagaggatg catagtctta tctcatctcc   72960
tagtcaaagt caacaccatg ataaataaga gtcaaatcct gagatgtgaa ttggggacat   73020
ttgagtggtt aaccctgaga agcttgcacc ttcagaccc tcaataccct tgctccccag   73080
agaaggctgg acattgacct cagcacaggc aggagccctg caagatgcca tttgtcctac   73140
taaagatgga ccctccact ctgtttctag gtaaataacc aaagtcaagt ctccacacag   73200
cctgagcaag aaagtcagag cctgctacag gagaaaatac cacactggcc aaaggattca   73260
ctagccctgg ccactgtgtg tgggaggaac caggggaatca tgtgtgggag tcaatgttga   73320
agctgtttga ctggggtgg ggtggaataat aagcctggcc ctggggagtt tttcccgttt   73380
gagggccttt acccacaact caagatccag tgctatagca ggagatcccca ggctagtccc   73440
taacagattg tcaggattga acttggccta gagtaaaatg aggaggatag tgccagaact   73500
ttctcaacat actattgagg aagaggtcag aaggcttaag gaggtagtgt aactggaaag   73560
gggtcctaat ccagaccca ggagagggtt cttggaccttt gcataagaaa gagttcgaga   73620
cgagtccacc cagtaaagtg aaagcaattt tattaaagaa gaaacagaaa aatggctact   73680
ccatagagca gcgacatggg ctgcttaact gagtgttctt atgattattt cttgattatta   73740
tgctaaacaa agggtggatt attttgtgagg tttccaggaa aggggcaggg atttcccaga   73800
actgatggat cccccactt ttagaccata tagagtaact tcctgacgtt gccatggcgt   73860
ttgtaaactg tcatggccct ggagggaatg tcttttagca tgttaatgta ttataatgtg   73920
tataatgagc agtgagggcg gccagaggtc gctttcatca ccatcttggt tttggtgggt   73980
tttggccggc ttcttatca catcctgttt tatgagcagg gtctttatga cctataactt   74040
ctcctgccga cctcctatct cctcctgtga ctaagaatgc agcctagcag gtctcagcct   74100
cattttacca tggagtcgct ctgattccaa tgcctctgac agcaggaatg ttggaattga   74160
attactatgc aagacctgag aagccattgg aggacacagc cttcattagg acactggcaat   74220
ctgtgacagg ctgggtggtg gtaattgtct gttggccagt gtggactgtg ggagatgcta   74280
ctactgtaag atatgacaag gtttctcttc aaacaggctg atccgcttct tattctctaa   74340
ttccaagtac cacccccccgc ctttcttctc cttttccttc tttctgattt tactacatgc   74400
ccaggcatgc tacggccccca gctcacattc ctttcctttat ttaaaaatgg actgggggctg   74460
ggcgcggtgg ctcatgcctg taatcccagc actttgggag gccgaggcgg gcggatcatg   74520
aggtcaggag atcgagacca tcctggctaa cacggtgaaa ccccgtctct actaaaaatg   74580
caaaaaacatt agccaggcgt ggtgcaggt gcctgcagtc ccagcggctc aggaggctga   74640
ggcaggagaa tggcgtgaac ctgggaggcg gaggttgcaa tgagccgaga ttgtgccact   74700
gcactccagc ctgggtgaca gagcgagact ccgtctcaaa aaaaaaaaaa aaaaaaaaaa   74760
tagctgggca tggtggcgcg tgcctgtaat accagctact ctggaggctg aggcaggaga   74820
atcgcttgaa cccagtaggc ggaagttgca gtgagccgag atcttgacac tgcactccag   74880
cctggtgaca gagtgagaca ctgtctcaaa aaaaaaaaaa agaaaaaaaa agacagaaag   74940
aaagagcaca gacagagtca caggtatttg cagtaggaag ctgtcaggtt agagtgcacg   75000
```

                                          73

```
gaaatagaaa gtatatttta cacttacagc acatcttcgt ttgattagcc acatttaaaa   75060
tactgaatag caacgtgtgg ctatttagta ttcactaaaa tcttggacag tgcaagtcta   75120
aagaatcctt gatccgtccg gcatggtggc tcacgccttt aatcccagca ctttgggagg   75180
ccaaggtgga aggatcactt aaggtcagga gttcgagacc agcctggcca acatggtgaa   75240
acctcgtctc tactaataat acaaaaaaaa ttagccgggc atggtggtgc atgcctgtaa   75300
tcccaggtac ttgggaggct gaggcaggag aatagcttga atccagtagg cgctgcagtg   75360
agccgagatc atgccatgcc actactgcac tccagcctgg gcaacagagt gagactgtct   75420
caaaaaaaaa aaaaaaattg ttgggcgtgg tggctcacgc ctgtaatccc agacttggga   75480
gaggctgagg ggggtggatc acctgggttc tggagttcga gaccagcctg gccaacatgg   75540
tgaaacccca tctctactaa aaatacaaaa attagctggg cgtggtggtg gcaggttgca   75600
atctcagcta ctcagaggc tgaggcagga gaatttcttg aacccaggag gcagaggttg    75660
cagtgagcca agatcgcgcc tctgcactcc atcctgggtg gcagagacaa actatgtctc   75720
aaaaaaaaa aaaaaaaatac ttgattgtct ggacattctg caagaaatgg caagtgttcc    75780
tcaagacaca tacatgccag aaggtgagat ataaactcta ctaagattca gtggcctgcc   75840
acactggtga catttttaaa cctgctagat gtttgtgtag aaaaggattt aaccttgccc   75900
aaagaggggt ctggccttt ccccagcta ctgacataaa tctcttaaa ctcttgaaat      75960
atcattcctg atagaagtat tttgtttg actagggc ttgggccagc cagatagcaa       76020
caatgtgatc tgggttgggg gctttgaatc aggtggcatc agtgtgacct cctgagtggc   76080
tagagactag aatcaaccac atgggcagac aaccagcttc acatgatgga attccaataa   76140
agactttgga caacaaggct tgggtaaggc ttcctggttg gcaatgctct atactggggaa  76200
acccattcctg actccataag gagaggacaa ctggatattc tcatttggta cctccctggg  76260
ctttgcccta tgcattttc ccttgtctga ttattattat tattatgaga tggaatctcg   76320
ctctgtcacc caggctggag tgcagtggaa tgatctcaac tcactgcaac ctctgcctcc  76380
ccggttcaag cgattttcct gtctcggcct cccgagtagc tgggactaca gatgcataccc  76440
accacacccg gctaattttt ttgtatttt agtagagacg gggtttcacg ttagccagga   76500
tggtctcgat ctcctgacct catgttccgc ctgcctcggc ctctcaaagt gctaggaata  76560
catgtgtgac ccaccgcgcc cagccccctt ggctgattat taaagtgtat ccttgagctg  76620
tagtaaatta taaccgtgaa tataacagct tttagtgagt tttgtgagca ctctagcaa   76680
attatcaaac ctaaggatag ccttgggggac ccctgaactt gcagttggtg tcagaaataa 76740
gggtgctcat gtgtgtacca tgccctctaa ttttgtagtt aattaacttt cacaacttta  76800
ttattaccgc ttacactcaa tgtttattca catttatcca cataccactt attctagtgc  76860
cttgcatcaa agacttttcta tctcatgtac tttattctgc ttgaagtaaa tcctttagga  76920
tattctttt ttttttttaaa ctttgcacat acatactttt atttttatt tattttaat    77040
tttgttattt ttgtgggtac gtagtagata tatgtattta tggagtacat gagatgtttt  77100
gatacaggca tgcaatgtga aataagcaca tcatggagaa tggggtatcc atcctctcaa  77160
gcaatttatc cttcaagtta caaacaatcc aattacactc tttaagttat tttaaaatgt  77220
acatttaatt ttgtattgac tagagtcact ctgttgtgct atcaaatata attttttttt  77280
tttttgagac agagtctcac tcagtggcca agactgaaag tgcagtggca caatgtcggc  77340
tcacttcaat ctctgcctcc ctggttcaag cgaatctcct gcctcagcct cccacatagc  77400
tgggattaca ggcacacacc accatgcca gctaatttt atatttttt agtagagacg      77460
ggttttcgcc atgttggcca ggctggtctt gaactcctgg cctcaaatga tctgaccacc  77520
tcagcctccc aaagtgctag gattacaggc atgagccacc acacctggcc ccaggattca  77580
attcttttagt gaggtctgct ggtgacaatt ttttctttt ttttgagact gagtctcgct   77640
gttgtcagct tgggctggag tgcaatagca cgatctcagc tcactgcaac ctccacctcc  77700
cggattccag caattctcct gcctcagcct cccaagtagc tgggatta caggcaccca    77760
ccaccacacg cggctaattt ttgtattt agtagaaatg ggggttccacc gtgttggcca    77820
ggctggtctc gaactcctga cctcaggtga cctcacacc ttggcctccc aaagtgctgg   77880
gattacagac atgagccacc acgcacagcct aattttttt gatatataaa ttgttggttg  77940
tttgtgtcat ctttgaaata tatttttgat ggatataaaa ttgttggttg atagttatta  78000
tcattattat tattattttg agacaggtc tcactctgtt gcctatgctg gggtgtagta   78060
gcctccctag tagctgggac tacagatgca tgccaccata cccaactaat ttttctattt  78120
tttgtagaga tgaggtcttc caccagttcta cccaggtcag gctggagtga caggctgtgag  78180
aatccaccca ccttggcctt acaaagtgcc cacctttct tctaacctcct gagtctagc    78240
tagcatattg aatattttaat atgatctttc tggcatccac tgtaactgtt taaaaaatca  78300
gctgttttact tggcactcatt ttttttttt ttttttttga gacagatgt tgccctgtcg   78360
cccaggctgg agtgcagtgg cgtgatcttg gctcactgca agctctgcct cccgggttca  78420
cgccattctc ctgcctcagc ctcccgagta gctgggacta aaggcgcccg ccaccacgcc   78480
cggctgattt ttttgtatt ttcgtagagt tgggggtttca ccgtgttagc caggatggtc   78540
tcgatctcct gacctcgtga tctgtccgcc tcggcctccc aaagtgctgg gattataggc   78600
gtgagccacc gcgcccagcc tctttttttt tttttttag acggagtctt actctgtcat   78660
ctaggctggt gtacagtggc gtgatctcag ctcactgcaa cctccacctc ctgcctcagc   78720
ctgcaaata gctgggatta caggtgtgca ccatcacgcc cggctaattt ttgtatttta    78780
agtagagatg gggtttcacc atgttagaca ggctggtctc gaactcctga cctcaagtga   78840
tctgctgcc ccagcctccc aaagattaca ggcatgagcc accgcacccg gccaagtagc    78900
actccttga aggtaatctg cttccctac ccctagcaat ttttaacaat ttttcttcat     78960
                                                                     79020
```

74

```
ttttatttcc tgaagttttg ttattaataa tctgtgtgca gatttctttg tatttctttt   79080
gtttgcagtt catagtgatt cttgaattag tgtgttggtt tctgttatca ccacaggaaa   79140
attgtcagcc gttagctttt caaatatttc cttgctaaat tctctcttct cccctttcgg   79200
tacaattgat ttgattaaaa ctaaaaccag ggccgggtgc agtgactcat gcctgtaatc   79260
ccaacacttt gagaggctga ggcaggtgga tcacctaagc tcaggagttc aagaccagcc   79320
tggccaatat ggtgaaaccc cgtctctact aaaaatacaa aaattaccag gcatggtggc   79380
acacatttgt agtcaggagg ctgaggcagg agaattgctt gaatccagga ggtggaggtt   79440
gcagtgagct gagatcccac cactgcagtc tggcctgggc gacagagtga gatgagaatc   79500
tgtctcgaaa aaaaaagtta tgaatgtttg ataaactata tttgttagaa tgtttgttgt   79560
agaatactat tcattgattt ttaaacaatg ttagattaaa ccattcactg gatttgtgat   79620
aattaactta ctgattttac ctcactgatt tgttgtaatt aatacaactg gtataaaaag   79680
actgtgacga ggccgggcat ggtggctccc gcctataatc ccagcacttt gggaggccga   79740
ggcaggcgga tcacctgagg tcaggagttc aagaccagcc tgaccaacat ggtgaaaccc   79800
catctttact aaaaatacaa aattagccgg tcgtggtggt gcatgcctgt aatcccagct   79860
cttcggggag ctgtggcagg agaatcactt gaacccggga ggtggaggtt gcagtgagcc   79920
gatatcgcgc cattgcactc cagcctgggc aacaagacga aaactccgtc taaaaaaaaa   79980
aaagaaaaaa aacacataaa acaaaacaac actgtgacgg ttcccaaaaa ttaggagcat   80040
aattaaagga actcctgata aaaattaatt ttatcttaca tgtaaactaa aatgacttta   80100
tgaagttaat tcagaaatac aatgcagggt attagtttgc cacagctgcg tattcagcct   80160
aatgtaattat tcttgttatt tttaaattct tctttttaact ttactccatat gtggatcatc   80220
aaattcaaa agattaaatg acaatactct tagcagcaag cttccctaag catataaaca   80280
tttttaatggg tgatgattca gaaggtaccc gaagaatatg tactgccaga tatcattcac   80340
ccccatatac ctgccgcaca gacatcccat tttgggaccc tggataaatg tgtgggtgga   80400
gagaaagata ggagaaagtg gtataagcaa atggctttgg agtctgattg acagcgattg   80460
aaatcctgtc tctacctctt aacagcctca tgatcctaca taagttaccc cgatcctcag   80520
ggccacatct gtaaattggg ggttgcgatg gcagccatct cacagggtct cttttcgggg   80580
aaaggcagga attatggatt aagtgagtca gtaattgtaa agcacttaat acaaggaggg   80640
cgcataataa gtacttcata aataatgacg gccattatca tgactgaggt gtatgcagct   80700
gtcggggatt acggcgactt cagaatttct ggtgggcagg gctcaaaggc agacaaatcac   80760
actggaagtc gaggtgaggc actgcttctg cacagactgc ttagctggag agaatgagga   80820
aggcttagag gagatttaga gagatttaga gtcctccgcc tccaactctg tgggatctgc   80880
tcccgtgtcca gagacattca ggggatttct cgcactctcc cctccccctac gtccctcccg   80940
ccccatccaa ctaaccacac aacacataca aaatagcccc tgcgaggttc tgcacgaccg   81000
aagggaacag gagaaaggcg ctgcgctttc ttgctgatgc cctgtacttg ggccctggt   81060
agacacagcc acttgtcccc tcagcctgca gagaaatccc acgtagaccg cgccgggtc   81120
cttggcttca gccaatctcc ctttggtggg ggtgggatgc acgatccaag gttttattgg   81180
ctacagacag cggggtgtgg tccgccaaga acacagattg gctcccgagg gcatctcgga   81240
tccctggtgg ggcgccgctc agcctcccgg tgcaggcccg gccgaggcca ggaggaagcg   81300
gccagaccgc gtccattcgg cgcagctca ctccggacgt ccggagcctc tgccagcgct   81360
gcttccgtcc agtgcgcctg gacgcgctgt ccttaactgg agaaaggctt cacctttgaaa   81420
tccaggcttc atccctagtt agcgtgtgac cttgagcagt tgactttatt tttccagtgcc   81480
tagttttcca gataccagga ctgactccaa ggactattac tcatctggag ggtttagcac   81540
agtaccgtcg catagtaaat ttccatgtca gttttggtta cctttcatgc acttgcaaac   81600
atgccatgct ctgaaacgaa ataggcacat ctttttttt tttttttta aggagtctc   81660
ctctcgccca ggctggagtg cagtggcgcg atcttggctc actgcaacct ccacctcccg   81720
tgttcgagat tctcctgcct cagcctcctg attagctggg actacaggcg tgccacgacg   81780
cccagttaat ttttgtattt ttagtagaga cggggtttgc ccatcttggc caggctggtc   81840
taactcctga cctcaggtga tctgactgcc tcagcctctc aaagtgttgg gattacaggc   81900
ataagccact gcatctggcc agaaatgaaa taagtaaatc ttttaacctg ctctaacaat   81960
atagtgaaaa gaccatatta ttattagagc aggttaaggg atttgcctat ttcgggttct   82020
agttatagtc ttaaacttgg acattccttgt agaaagtaaa aagtttcctc ttcaaagttc   82080
ccttcttgt taaagaatac atcataagtg ttagaagtaa tagtttattt taaagactaa   82140
ctttcttcaa gcctccttgc tttgtgctaa taactctttg ttaagcccta tcctatgtaa   82200
ctgttgacaa tgctcacagg cacgttccag ttcacagcct atgccccttc cttatttgga   82260
aatgttattg cttccttaaa cctttcggta agcaacttcc tctccttctt cgttcttcct   82320
tgcacttacc tatttagaaa gtttttaggct attagcaaat cggctatcag tttaagagtg   82380
tgaggtcccg ctccagccaa tggatgcagg acatagcagt gaggacgacc caaatgcgta   82440
agggataaat atgtttgctt ttcctttgtt caggtgtgct ctcgacatcg ttccatctgc   82500
gattgagcac cctttctgca gaaagtaaag attgccttgc tggagatctt ttgtctccgt   82560
gctgactttt cttcgtgtga ccgattatct attttctaaca attttggtat ttctaacatt   82620
ctgaacaatc ttgggctagt tgtctccttct gggcctgttt cccatccgt cacatgataa   82680
acttcattgg tttaaaaacc ccagcgaaca tttattgagt tactattacc ttcggggttct   82740
ccccaacccc aaccccaggg agcagttaca acctcagccg ctgacgcgac tcgccggggtg   82800
ttaagaagca ccaaagacag ggaggcttga ttgattttgc ttttgggagta gaaaggcaga   82860
agattcacag gaaaatggca tttgagcaag gatgattcac tggagctagc tttaaaatac   82920
tggcgaggct tttatgttgc agtccttac aaagttgagc attcgcaggg actgcactcc   82980
gaaataagcc cgcttcccct tttcattcgc taatgatcca gggagctgct ggttccggcat   83040
```

```
gcggcaggtt gtgccttttc ctaatcaggg ttctgcatcg cctcgaaccc gcaggccgtg 83100
gcgggttctc ctgaggaagc agggactggg gtgcagggtg aagctgctcg tgccggccag 83160
cgcctgtgag caaaactcaa acggaggagc aggaggggtc gagctggagc gtggcagggt 83220
tgaccctgcc ttttagaagg gcacaatttg aaggggtacc aggggccgga agccggggac 83280
ctaaggcccg ccccgttcca gctgctggga gggctcccgc cccaggagt tagttttgca 83340
gagactgggt ctgcagcgct ccaccggggg ccggcgacaa acgccacaaa acagctgcag 83400
gaacggtggc tcgctccagg cacccagggc ccggaaaaga gggcggggta gcacgcgcgg 83460
gtcacgtggg cgatgcgggc gtgcgcccct gcacccgcgg gaggggatg gggaaaaggg 83520
gcggggccgg cgcttgacct cccgtgaagc ctagcgcggg gaaggaccgg aactccgggc 83580
gggcggcttg ttgataatat ggcggctgga gctgcctggg catcccgagg aggaggagg 83640
gcccactccc ggaagaaggg tcccttttcg cgctagtgca gcggcccctc tggaccccgga 83700
agtccgggcc ggttgctgaa tgaggggagc cgggccctcc ccgcgccagt ccccgggccac 83760
cctccgtccc gacccgggcc ccgccatgtc cttcttccgg cggaaaggta gctgaggggg 83820
cgccggcggg gagtcaggcc gggcctcagg ggcggcggtg gggcaggtgg gcctgcgagg 83880
gctttcccca aggcggcagc aaggccttca gcgagcctcg acctcggcgc agatgccccc 83940
tgagtgcctt gctctgctcc gggactcttc tgggagggag aaggtggcct tcttgcgcga 84000
ggtcagagga gtattgtcgc gctggttcag aagcgattgc taaagcccat agaagttcct 84060
gcctgtttgg ttaagaacag ttcttaggtg ggggttagtt tttttgtgtt tctttgagga 84120
ccgtgatca agatcaagga aatctctttta gaacctttatt atggaagtct gaagtttcca 84180
aatgttgagg gtttatgtc taaaagcaac acgtgaaaaa attgttttct tcacccagtg 84240
ctgtcttcca atttccctctt tgggggaggg ggtagttact gctgttacta aaataaaatt 84300
acttattgct aaagttcccc aacaggaaga ccactacttt tgatgacttt ggcaagtttg 84360
ctaactactg gaaccctaac ttacaaacga actacttaca ttttttgattt ccagttgtaat 84420
tacctgccca atgtttacgt agaaacagct taatttttgat tctgggtaac gttgttgcac 84480
ttcattaaaa atacatatcc gaagtgagca agtatgggtc tgtggacagc agtgattttt 84540
cctgtcaatt cctgttgctt cagataaaat gtaccagaca gaggccgggc gcggtggctc 84600
acgcctgtaa tcccagcact ttgggaggct tggcgggtgg atcacctgag atcgggagtt 84660
caagaccagc ctgaccaaca tggagaaacc ccgtgtctac taaaaatca aaattagcca 84720
gggtggtggc gcatgcctgt aatgccagct acttgggagg ctgaagcagg agaatcgctt 84780
gaacctggga ggcggaggtt gcggtgagcc gagatagcac cattgcactc cagcctgggc 84840
aaaaagagcg aaactccgtc tcaaaaaaaa agtaccagac agaaatgggt tttgttttct 84900
tttttttgttt tgagacggag tttcgctctt gttgcccagg ctgagtgca atggcgcgat 84960
ctcagtctcg gctcactgca acctctgtct cccagggtta atcgattctc ctgcctcagc 85020
ctcccaagta gctgggatta cccatgcccc accatgcccg gctaatttt gtattttag 85080
tagaaacggg gcttcaccat gttaggccag tcttgaaccc ctgacctcaa gtgggcctcc 85140
cacctcggcc tcccaaagtg ccaggattac aggcatgagc caccgcggcc agccagaaat 85200
gggttttgga aaaagcacta aacaaaatcg aacttggttt catatgacag ctctgctgct 85260
aactgtaaca ggggcagacc agttaaccta cttttctgtc ttctgtcagc tgagaattag 85320
atgattccca aaggcccatt gaactctgaa tgactttaaa tacttcttct taagtgggta 85380
cacggttttg gtaactgatg ccaggtgatg aatgcatgaa agtgcttaat gaatgaaacc 85440
ggtaaaatag taggaggaag cttttattggt aaggcagggg tatacctaat agctctctaa 85500
tttattggta ttgaagtggt taactttttgt ttttttaagg ggggaaaaca ttctaagaat 85560
aatgaggcaa actgcatatt gcacaagaga ctgttgtctc tattcaacaa ataccttttg 85620
agtgtccaga gtctgccagg tgctgtgcta ggccctcacg attgagtagt gaaccagaga 85680
atgtccctgc acccatggag cttattgtct actggggtag acagtaata aataagcaaa 85740
caaatcttct ctcttctccc ttttgctcca tgtaagtgtg tgtgtatagg tgtactactta 85800
caagttgagt aaagtgttat gaaagattaa gaggagaaat gcattttggt tagatgttag 85860
aggactcagc aggtgacctt gaaacttaga gctgaaggat cagtaggagg taactagaga 85920
ggccaagggaa tcgcatgttc aaaggccagg aggcaagaaa gagcatggtg ccttcaaga 85980
gaggaaagaa ggctactgtg actggagcat agatgtaggc aagtgttggt tgattgagag 86040
ctctacgggc catggttagg ttttattcct aatgccgaga tgccaaacat ggtggttcat 86100
atctgtaatc ccagtatttt aggaggccga ggcaggaata tagctgaacc ctaggagttc 86160
aagaccagcc tgagcaacat gagacctgta caaaacattt aaaaaattgc tgggtatgat 86220
ggtgcacacc tgtggtccca gctactcagg aggctgaggc agaaggatca cttgagccta 86280
ggaggtggag gctacaatga gccatatttg agtcactaca ctccagcctg gatgacaaag 86340
tgagaccatg tgtcaaacaa aatacagaaa gaatattaat ttaaaatttt agaagaaattg 86400
tgatctgaac ttatatctta aaaagatcat tctagggcat ggtggctcat gcctgtaatc 86460
aagggctttg ggaggctgag acaggaggat cacctgaggc cagttcgaga tcaacctgga 86520
cagcatagag agactccatc tctacaaaaa gaaaaaataa atagctgggt gttgtgagtt 86580
attcaggagg ctgaagcaga aagatcactt gagcccagga gtttgaggct gcagtgaagct 86640
atgatcccac cactgcaaca cagtgagatc ttgtctcaaa aaaaaaaaaa aatcattcta 86700
ggtgctttt ggaggctgga tgtggtaaga gtagaagctg gagatggtcc tgttagggat 86760
tcgattcaga ctttaaatac catcaatgca ttgagtccca aatttacatc actacgttgg 86820
atccttgccc ctgaatccag actggtatat ccaactttag gttcagtttg tatctctacc 86880
tgaccaatat agaggtgtcc agtcttttgg cttccctagg ccacattgga agaagaattg 86940
tcttgagcca cacatagagt acactaacgc taacaatagc agatgagtca aaaaaaaatc 87000
gcaaaactta taatgtttta agaaagttta cgaatttgtg ttgggcacat tcagagccat 87060
```

76

```
cctgggccgc gggatggaca agcttaatcc agtagatacc ttcaacttac aatatctaaa    87120
attttatgcc agatttagtc attttaaacc tgctcatcag tttttctcaa gaagtagtat    87180
tttggctttt tttcttttct ttttttttgag atggagtttc gctcttatcg ttcaagctgg    87240
agtgcagtgg cggatcttgg ctcactgcaa cctccgcctc ctgggttcaa gcgattctcc    87300
tgcctcagcc tcgcaagtag ctggaattac aggcatgcgc caccatgacc agctaatttt    87360
tggagacagg gtttcaccat gttggtcagg ctggttttgt actcctgacc tcaggtgatc    87420
tgcctgcctc ggcctcccaa aggctgggat tacaggcatg agccaccgct cccggctgca    87480
tttttggatt tttagttgct cagcccaaaa ctttagtaca tctttgaacc tcttctttcc    87540
tcctactcta tatctgatcc atcagcaaat ctgttaggtc tacctcacac atatcgaaat    87600
cctaccacgt ctcaccatct gtgacaatta acacctggt ctaggcagtc atctctgtta    87660
agattgagtg gttaaggatg tcctctaagg agatgacatt caaatcttag cttaaatgtc    87720
aagagggagc tggtttata aagattgagg aggcagcatt attttgccat aggcttccat    87780
ttggtttcca ttccattctt gatacttatg gtatatattc aaaacaaatg cacagaaaca    87840
gacccaggta tattgggaat ttcggatata gagttcctag ttgggaaaag atagactgat    87900
ctgtaaatga tgctagttat ccatcatctg gcaaaaata atttcctgcc tcctctcata    87960
tatctcagat caacagactt tttctgttaa gggccaaatc ataaatattt taggctttcc    88020
agaccatatg gtttctgtca cactctcctt tatccttgaa gccatagaca atatgtaaac    88080
aaatgggcat ggctgtgctca cgataaaact ttacttacaa aaactggtag tgggccagtt    88140
taggcatggc cagcactttg ggaggctaag gcagatggat cacttgggtc caggagtttg    88200
agaccagcct ggccaacatg gtgaaaccct gtctctacta aaaatacaaa aaatagctgg    88260
gcatggtggt ggtgtctcat aatccagct acctggaggg ctaagacagag agaatcactt    88320
gaacccagga ggcagaggtt gcagtgagcc gagatagcac cactgcactc cagccagggt    88380
gacggagtct taaagcaaaa caaaaacaaa ggtagtgggt tgtatttggc ccatgggctg    88440
tagtttgcca atccctgatg cagaaacaaa ttccaggtaa ataagagcct ggaatgttaa    88500
aaaaacaaaa cttgaagtca tgtagaagaa caggtagggg gaacaatcct gatctcagga    88560
taggaaggga tattgcttaa aataagacac aggaaaatat aatccatgtt gtgtaaattt    88620
gactacgtta aaacttaaaa ctttcgccaa gcgcggtggc tcacgcctgt aataccagta    88680
ctttgggagg ccgaggtgag cagatcacca ggtcaggaga ttgagaccat cctggctaac    88740
acggtgaaac cccgtctcta ctaaaaatac aaaaacattag ccgggcgtgg tgggcgggcgc    88800
ctgtagtccc agctacttgg gaggctgagg caggagaatg gcctgaaccc gggagggcgaa    88860
gcttgcagtg agctgagatc gcgccactgc actccagcct gggcgacaga gtgagattcc    88920
gtctcaaaaa aacaaaacaa aacaaagcaa aaaacctaaa actttcatac aataaagtat    88980
acctaagata cttctagaag agaagattta catccaggac gtgtatggaa tttctgcaag    89040
taataagtaa aagacaaggg acatgaagag gcagttcaca aaagaggaag ccaaaatgac    89100
caataaacat gaaaggatgt ttaacctcaa aggaaacaag gaaatgaatt aaaaacatca    89160
aatgccattt caaaactagt aagttggcaa aattaaaaat accaaggatg agaatatgaa    89220
gcatggctat atgagtgcat ggaatggtac agtcactttc attaaaaatg cacataatttt    89280
gtttttattt tatttttttg agacagtcta ttgtcgcccag gctagaatgc agtggcatga    89340
tctcggctca ccacaatctc tgcctcctgg gttcaagcaa ttctcctgcc tcagcctcct    89400
gagtagctgg gattacaggc acatgccaca acgcccggtt aagtttttgta tttttagtag    89460
agacagggtt ttgccatgtt ggccaggctg gtctcgaact cctgacctca ggtgagctgc    89520
ttcccaaagt gctgggatta gaggcgtgag ccaatgctcc tggctgaaaa aaatgcacat    89580
aatttgttac ctagcaattc catgtctaga ggcttatcct agagaaattc ttgcttatat    89640
gcataggaag acgtgtacta gaatgttcac tagttgaatg tttaagtgaa aattaaggaaa    89700
taaagtaaat gttcattaac aggaaaatga gtaaaggtat atttataaa caattaagta    89760
gctaaaatga ataaatcagg gctgcgtgaa tgactagaaa ctggtttcaat agtcatgtca    89820
gattattgaa tgaatacagg tcagatatgt atagagtgtc atttgtgtaa ttaatttttt    89880
tttttttttt gagatggagt ctcactctgt tgcccaggct ggagtgcagt ggcgtgatctt    89940
cagctcactg caacctccac ctcctgggtt aaagtgattc tcctgcctca gcctcccgag    90000
tagttgggat tacaggcgta caaccaccg ccagctaat tttctatttt ttagtggcca    90060
cagggtttca ccatgttggc caggctggtc ttgaactcct gacctcaagt gttccaccca    90120
acttggcctc ccaaagtgct aggattacag gcgtgagcca ccgtgctcag ccatttggcgt    90180
gattttaaa gatgtgcaga ataatgccat taaaaaaaat acacatacat gtatatatat    90240
acacgtttgg ctgggtgtgg tggctcacac ctgtaatccc agcactttgg gacggccgagg    90300
caggaggatc acttgagccc aggtgtacaa gactagcctg ggcgagataa caagacccca    90360
tctcaacaac agaaagaata attaggtatg gtggcatgag aggatcactt gagcccagga    90420
gttcgagtgt tatcaggcca ctgcactcta gcctggacaa caaagcaaga ccgtgtctca    90480
aaaaaataaa aataaaaagt atttgtatgt ggtcatagtc aaaaaacgta catggaagga    90540
aaatgtcttt atttatttat ttattttttt ttttttaaga cagatcttg ctctgtcacc    90600
caggctgggg tacagtggtg taatctcagc tcaccgcaat ccggggttcaag cgggttcaag    90660
cgattcttct gcctcagcct tctaagtagc tgggactaca ggtacccgcc accacaccct    90720
gctaattctt gtgttttcag tagagacagg gtttcaccat gttggccagg ctggtctcga    90780
actcctgacc ttaagtgagc caccgcctt ggctcccaaa agtcctggga ttacaggtgt    90840
gagccactgc gcttggccag gaaatatcta atttagtaag tatttatatc tgggaaaagga    90900
agggtcaggt ggtgattcat aggaactcta aagtctatgt ataatactta gggggacaga    90960
aggaaataaa gcaaaatgct gatatttgat tgttgagttg tgtatatgtt agaagtataa    91020
cataggagat ctgattgata gtaggagaat gttttttaggt ggtaaaagtg gaaccgtggt    91080
```

77

```
ggtttgtttt ggcagtagaa tcagttggtc atagtttgta tgtggaaggt aataaacaga    91140
ccatgttaag gatgacttcc ggaattttgg tctgagtagt gggtggatga cagtgtcatt    91200
catgcagggaa gatgaagact gaggtaggaa caggtttggg agaagatgac atgttccctt    91260
ttagacaagt ggaattatgg aagatggcag gtaggtggtt agctatatga atttggatga    91320
aaagatttag gatggagata taaatttagg agtaacagcg tatctatggt attgtaagcc    91380
ttaagaatgg gtaggatcag ccaggaaata cagatgtata tgcagaagag aggagtcaag    91440
gaagccaaga caagttaatg tttaaagtga gtgatgtagt ccatgggcag atgctgctga    91500
gagggctgca aacaccagtg accctacaac attttttaaat gtcgtcttcc tgacagcagt    91560
gatcagtacc tgcaacgatc ttatttattt ttttcatgtt agtctccaca cacttgaatg    91620
tagacttttt gaaggcaaaa tcattgcctt ttctgagctg ggagcatgtc tggcacatac    91680
caagcactca acagttgatg tattgacttc atccagatac tctgagggcg agttatttcc    91740
tgctactagc ctttcacctt tcaatgttta agagcacaaa tacagagatg gcacatggtt    91800
ggcatttgtt attttgataa ccttttcctg gtaagatttt ttaatgttga aaaaaaaaaa    91860
caagaaaaga gggttaaaaa tagtcttatg tcagatcctg tgatagaatt cacacttggc    91920
ttaagctgct gggcaccttc ctatcttgga tgtcatatta gcttatctac agcagaattt    91980
ttactgtttt atgtagtaag gaagcaatta tatgattatt ttacagacaa attattcttt    92040
atctttatt ttttttagacg gagtctctct ttgtctccca ggctggagta cagtgtcgcg    92100
atctcggctc actgcaacct ccgcctcctg ggttcaagca attctctgcc tcagcctccc    92160
aagtagctgg gcttacaggt gtccgccacc acacccagct cattgttttg tattttttagt    92220
agagatgggg tttcaccatg ttggccaggc tggtcttgag ctactgacct caggtgatcc    92280
acccgccttg gcatcccaaa gtgctggaat tacaggcgtg agccaccgtg cctggccag    92340
acaaattatt atactctgag tgttagaggc ttaggatgtt tcaattgat gctatgggag    92400
gaataagtaa taagatatga tacacaacca aagacttttc ttcactatgc ttctagtagc    92460
tagtactatg gatgacacat ggtaataata ttggttagca tttgtcctca atttactgtg    92520
ctagttactc ttctaagccc cttacaggta tatatttttt ttcatcaata atcctctaag    92580
gtagtttta ttattgacct aattttataa atcaagaaaa ttaagacccca gagaagtaag    92640
taacttgtcc aagatcacat ggcttataag tggtagagcc agaatttgac cccagatgtt    92700
gtgactacat tgtctctcca taagcaggtt caactctttt gactggatgc tgttccaagg    92760
tcacttcctt agagaagcct ttgctgacaa ctaccctcct gtgccctcct ccaaggctgt    92820
ccattgttct agaacttttga atactcatct tagaataaag ctggtctaat ttttacagtg    92880
ttatagaatg gatctctgac tgcaaaagtt ggtcataatt atcttttttat gttctagtga    92940
aaggcaaaga acaagagaag acctcagatg tgaagtccat taaaggtaag ttctgccctt    93000
ggcagtccac tgcattaaaa agtgatgtgc tttgcatttg tgagttcttt aatcctgtta    93060
tactctctct tttggcatta atcattctg ccttattttа taattactta tgattttgat    93120
ttatttccct ctttaacctg tataatgctt taacatctag catataataa gtaggcttttt    93180
tttttttttt ttttttttgga gacggagtct tgctctgtta cccaggctgg agtgcagtgg    93240
cgcgatcttg gctcactgca agctctgtct cccgggttca caccattctc ctgcctcagc    93300
ctccccagca gctgggacta caggtgcacg gcgccacgcc tggctaattt tttgtattttt    93360
ttagtagaga cagagtttca ccatgttagc cagtatggtc tcgatctcct gacctctgtga    93420
tccgccgcc tcggcctccc aaagtgctgg gattacaagc gtgagccacc gcacccggcc    93480
gtaagtaggc tttttttacc ttaattttat tttttttggag tggagtcttg ctcttatccc    93540
caggctggag tgcagtggtg ccatctcggc tcactgcagc atccacctcc cgggttcaag    93600
cgattctcct gcctcagcct cccgagtagc tgggattaca ggtggccgcc accatgccca    93660
gctaatttt gtatttttag tagagacagg gtttcaccgt gttggccagg ccagtctcaa    93720
actcctgacc tcaagtgatc cactcgcctt ggcctcccaa agtcctgaga tacagacgt    93780
gagccacca gcctgcccat aagtagcctt ttactgagcc ttgtgtgtat tggctatcct    93840
agtgattaca gtgaaccagt gcccttctta ttaatcacac atttaattgt tccctaaaag    93900
tgattagttc actttattta tttagtaaga caaaaatga agaatactct taactgagca    93960
gtctgttcac tgtaggaaag cactgacact tataaggctt agttttctgt catttatcca    94020
gaagtatggt tgattacagt ttttacttt ttatttgaat gaacaacctt aatttaaaat    94080
atatttttgt tatttttttgt tgggatcgat acattgtcct tgtttataga ttagagcatg    94140
cttttttaaag atgctgtatt actcactgat tttatttgtc cagtgtacag agattgaagt    94200
gggaaaattaa taatggaaat tgtttccata gtcattacat attaatttca tcaatttatt    94260
tccataaaat ctgtagattg ctactttttt agattttttcc ttcaaatgtt tttatgttgt    94320
attgcttgca ctgatgtattt attctctatg ctcaatttgc tggagaagaa gactaattat    94380
aacttaggca agttgtaaaa ttagggaaaa aagtaaggta ccttacagcc tagtttactt    94440
atttctttatg taaagccagt tagattccac attagttcaa actgccttct ttgagcaaaa    94500
cttgattggc agtgataaag gcttaaagcc cttctcaagc agagacctgt aaagactaga    94560
tctgactgta gtagaaggaa ggaacttaga tgtttcaggc agtgagaaca ccagtcttcc    94620
actctaaact ttgccactaa cagtatgacc ttgggaagtt gtaactttct tcagattcctt    94680
catttgttga atggggggat tggcctagct aatttctaaa tctctactgg gctaaaaaat    94740
tctgtgctta tactctgatt atgaagtaca taatctgtgc ttaacattca ctgacttatc    94800
cttaggataa tacagaagca gtacaagaaa cagccccctca agatgtttgc agtctggtta    94860
gaaagacaaa cttatacaca gaacagtagc aaatagcaac aaataataat agctgccatt    94920
tatagaacac ttcctctgtt ctgggcatta gacaaaaact gactataacg gtgaacaaaa    94980
aagacttagg tcctgccctc attgaactta cagattagta gggagaggga acattaatca    95040
agtaattcca cagatggctt agcctagatt ggtagtgatg gaagtaaaga gatgtgaacg    95100
```

78

```
gacttgaaaa aaaattcgga ggcaaaatgg atagaagttt attattgatt aaatatgagg    95160
tgtgagagag agggatattt aagattgata cctaccttct ggcttgccta acagaaccaa    95220
aacaggaaat tatatgttca gttttgttat gttgggtggg aggtgctttt gagtcattca    95280
tttatatatg ttatatatgt tattttatat gcatagtaat tttaaggtct gagttttaaa    95340
ccaaaggtta gagagtgatt ttttagagtc tagcaaacct aagttgaaat cctgcctgtt    95400
gaaatggctg tttactagct cattaaccta gggcaaagta ttcaacttgt tttcattttt    95460
gtcttcatct ctaaaatgag gaaaatatgg tcttacaaga ttgtcctgag agatagatga    95520
aataatatcc aaaaaaaaaa aaggtacata gagaaactcg tatagtgcct ggtatatagt    95580
aggtcctcca ttggtagcta tcattatcta gttttaacat agccttcagt ttgttgaatt    95640
agtcaaactg agtgaagcac tgcaaggaat tcagaggaat ttgagatcaa caaatgatgt    95700
ctgaagttta gggaagactt catggcaatg acacttacct tgtataaaag ttgaagaata    95760
agaaagattt gaatgagaga ttctttctct tctccctacc agcccagctt ctttattgag    95820
gatatattgg gcaaaggggc cttcagacaa gtagagggag atttttacag aaagattgag    95880
atgaaggtat agaaggcctgt aaagaccaga aaagagaatt gagacagagg aagcaggaag    95940
ccactgtagg tttttgagca agatattgat gctgtaagta tggtgtttat gaaaggttag    96000
tctggaagag atttgcagga tggagacccc ggaagttttt ttgttataat acagaaagac    96060
ttgcactgag ggtgaggtgt taaaaataaa caggtaagta aatgtttaaa catcttgaag    96120
gaaaagtcaa caaatcttgg caagtaaaca gataacagtg aaaaagaatg ggaccaagat    96180
tttgagtttt ggagactggt ggattgaaca gacagggaaa ttgagaggag aatcagatga    96240
tgatgttttt agttgatatt tagacagatt gtgcttgaga tggtaaagtc aatgtgggtg    96300
ggaatgctta gtagcgagta atcagtgata caagaccaaa gcccaggtca aagacaagtc    96360
acagatacag atcagggctt tttcatctgc tccacagagg tgtacccctag gagctgttgc    96420
aaacagtcca tgtggagggt gtgagtaaga tgtttccctt gaatttgcca gaattactttt    96480
tttgttgttg ttgttgtttt ttctgagaca gattctcgct ctgttgccca ggctggaggg    96540
cagtggcgag atcgcgcagc tcactgcaac ctctgcctct cgggttcgag tgattctcct    96600
gcctcagcct cccaagtagc tgggattaca ggcttgtgcc accaagccca gctaatttct    96660
tttgtattt tagtagagac ggggtttcac catgttggcc agactggtct cgaactcctg    96720
gcctcgtgat ctgcctgcct cagcctccaa aagttctggg attacaggcg tgaaccactg    96780
cacccggtcc cttgttaagt ttatttttggt gggaagcaaa ggaggtttca gctttttaaaa    96840
agtttgaaaa ttattgctct ggtaataatt aaagatttga gagtaaatat gctttctagc    96900
agaaagaata aaagaagaac agatagcctc aagaaggggga gccaaagaag caggctatat    96960
ctgacacact gggtgttgat aaatgggtat taaaagaatg agagcaatga gcagatagaa    97020
gaggaaatta ggagagtata ataccatgga gaccaagaaa gatagactat caggagggag    97080
tggtaaaaat aagttactag ttctaagaga gatgttaaga gggaccgggg aaagccttgt    97140
acaaatgagt tagtagcatt ttacattata tacatctaat taagaaacaa tgcgagcatgt    97200
tcaccattcc tatagactct tacttgtact tgtctgaaca cgaaaactgg cttttgtttta    97260
taaataagct aaaaattatt ttgctccaat ttctcatgaa aataaaaata aacctttcttt    97320
taacattgaa aaaatagttt gaagacagtc actcttcatt ttgtaattcc cacaactatt    97380
attgaatgac tgaaattatc tttattctga agccaaaggg gtgatactga tatttcttca    97440
gactactaaa aatatatttt atgaattttt agtgtgcttt atcttttttt gttttttttt    97500
ttgagatgga gtttcactcc cgttgctcag gctggagggc agtggtgcaa tatttcagca    97560
ctgcaacctt cgcctcccag attcaagcaa ttctcctgcc tcggtctccc aagtagctgg    97620
gattacaggc acctgcccc acacccagct aattttttgt atttttagta gagacagggt    97680
ttcaccatgt tggtcaggct ggtcttgaac tcctgacctc aggtgatcca cccaccttgg    97740
cctcccaaag tactgcgatt gcaggcatga gccaccatgc ctggcctgag gaattatttt    97800
ctaggttccc cccaccccaa gcatttattc tgcaattta gttttgttcc taaagcaagc    97860
aaggtttaag gatttaaaaa taatccgtat tttagaatgc tttctggctt tgttactttt    97920
tatccacagt agaagttctc agagaatgat ctccctcttt taatttaact ttttggcaca    97980
gtattttgag aattataaat aatattagaa tgttttctgg ctgggtgtgg tggctcatgc    98040
ctgtaatccc ggctacttgg gaggctgagg caggagaatc acttgaacat gggaggcaga    98100
ggttgcagtg agccgaggtc atgccactgc actccagcct gggtgacaga gcaagactct    98160
gtctgggaaa aaaaaaaaaa aaaaaaagag tgttttcttt cctattttcc accacttgat    98220
taagttactt ttctcttaa gtattttttg ctgagtatgc tgacttaaga gtaatgttac    98280
aaaatttaat ttttaaagtt ctctgaaagc cccttttatga gagttttagg ctatggaaat    98340
gtgtttaatt cttaacaatt ttttgaaaaa ttatagcttc aatatccgta cattccccac    98400
aaaaaagcac taaaaatcat gccttgctgg aggctgcagg accaagtcat gttgcaataca    98460
atgccatttc tgccaacatg gactcctttt caagtagcag gacagccaca cttaagaagc    98520
agccaagcca catggaggcc gctcattttg gtgacctggg taagtaacta tcattttata    98580
ttaacttgta ttagaaggat ttgagtacaa tatgtgaaac ttctgtcata ggatacagaa    98640
ctatataatt ggaaagtgct ttggaaaaaa tgtatttaaa ataacagcta caagtataat    98700
gggtagctgt gttgtgttcc tgtaaatata gaatataaag catgcccagt agaaaaacaa    98760
gcatttccag aagaaaatata tctgatcact aaatataaat atatgaaaaa gatgtctcac    98820
tttattactg agggaagtgc aaattaaaat aatcagttaa tgttctccta acacattagc    98880
atatttttta aagtttgaca atttgaatgt cagtgaagat gcagggaaat accctcccta    98940
tttagtgata atataatctg gtgaagactc tttggaaagc aatttggaaa tcagtataaa    99000
atatgcatgt catttaggcc actctttcta agacctagcc ctcagatatg ctcattcata    99060
tgtgcaggtg tgtatgtgtg tgtgtgtgtg tgtgtgtgtg tgtatatgta tgtatgtatg    99120
```

79

```
tatgtatgta tgtatgttga aggctattca ttatagtatt gtttgtgata gcaaaaaatt    99180
atggacaaca tataaatatc tgttataggg aaataaccaa attgtggtat acgcatgctc    99240
tggagtataa tatagccatt tgtttctatt tatttatttt cttgagacag ggttttactc    99300
tgttgcccag gctggagtgc agtggtatga tcatggttca ctgcagcctt cacctcctgg    99360
gcacaagcca ttctctcgcc tcagcctcca gagttactag gactgcaggc atgtgtcacc    99420
acacccagat aatttttaa ttttttgtag agacagggtc tcactatgtt gcctaagctg    99480
gtctcaaact cctggcctca agcaattctc ccacacaggc ctcccaaagt gctgggatta    99540
ccaacgtgaa ccaccacacc tggttcagtg tagccattta gaaatctaaa aaagacgtgg    99600
gaaaatgtct aaggcatgtt taaatgtgag aaaagcaagt cacagtatgc atggtaaaat    99660
ccgttatatt aaaataagtt cttccaaaac aaaaacatat gcaggagacc tttattttgt    99720
cagtatttct tacccaaatt tctgcactta gaaaattgca tgtcatgttg tcataagttg    99780
aaaaaaagat ccatgaacca atggacttct aataaaatca gtcctgcttt tgacatctct    99840
ctctactttt gtgtatattc aaaccagagt gtcaatgtgt ttgtggggca cacttagcaa    99900
taatacatag cagacaaaat gcatatagct cagagagtaa aattgtaagt tttgctagat    99960
cactcataaa ttgctgatga gaatttaaaa tggtgcagat gctctggaaa acaggcagtt   100020
tctttctttt ttttttttt tcttttttgag acagggtctc actctgttgc ggcaggctgga   100080
gtacagtggc gtgattacaa ctcactgcag cctcacccte ctcaggttca ggtgatcctc   100140
cctcagtctc ctgagtagct gggactatag gcatgcacca ccacgcctgg ctaatttttg   100200
tatttttttt tttttttttt gtagagacgg ggtttcgcca tgttcccag gctggtctca   100260
aactcctgga atcaagcgat ccacttgcgt aggcctccca aagtgctggg attacgggcg   100320
tgagctactg tgcctggcct aggcagttg tttgtttgtt tgtttgtttg tttatttatt   100380
tgtagacgga gtctcacagg ctggagtgca gtggcccaat ttttggctca ctgcaacctc   100440
cgcctcccag gttcaagtca ttctcctgcc tcagcctcct gagtagctgg gatgacaggt   100500
gcctgccata atgcctggct gattttttgta tatttagtag atatggggtt tcaccatgtt   100560
ggtcaggctg gtttttgaact cctgacctca ggtgatcagc ccgcctcggc ctcccaaagt   100620
gctgggatta caggcatgag ccgtcatccc tggctggtgg tttcttatga cgtgaaacat   100680
gcaattacca tatgacctag cagttgcact ctgtatttat cccagataaa tgaaaactta   100740
cctcccaata aaaacctgtg cacaaatgtt catagcagct taatattgaa aaactggatg   100800
ttcttcagca ggtgaatgaa ctggttcatt cataccatgg aataccatc agcacaataaaa   100860
aggaacaaac tgttgataca tttaaccacc tggatgaata tcaagggaat tatgctgtca   100920
gacaaaaacc agtccctaaa gactacatat agtatgattc cgtttggata atattcttga   100980
aatagagaaa ttaagagaaa tgaaaagatt agtgtttgcc agatgttaga gacagggagg   101040
tgagaggggt aagtgggtgt agttataaaa gtgcaacatg agggatcttt gtgatgttga   101100
agttgtatct tggcagtgga tgcagaaatc tcaatgtgat aaaattacaa agaactaaaa   101160
acaagaatga gtatagataa aactggggaa atctgaacaa gttagagtgt tgtatcactg   101220
tcagtatctt agagtgatat tgtactatag ctttgcaaga tgttaccatg ggagaaacta   101280
aagtgtacaa gggatctcta ggtattatta tttttttaga gatggggttt cactatgttc   101340
cccaggccgg tcttgaactc ctgggctcta gtgatccgcc tgccccagcc tcctaaagta   101400
ctggaattac aggcgtgagc gaccatgcct ggcccttca gtattgtatc ttagaacttc   101460
atgtgaatct agcattatct catagaattt aattaaaaga aattgtaaac ctcacagaag   101520
atcagaattt cctcaagttt gtgatgttga caaagatgaa ctagttgaca ctgacagtaa   101580
gactgaggat gaagacacga cgtgcttcaa aaaaatgatt tgaatatcaa tggattaaga   101640
agaactcttt tgacaaattg atgaaacct cagtccagttt tataagaatg ccccatcttta   101700
tgatcatgct atgaaagcca attttttaaa aaattttttg tcttttcctaa caattagctt   101760
gtggttataa ttaaaattta gttaaatata agataaatga ttttttatta agtttagttt   101820
cattttttcaa ggtacgatct caaagctact cttaaccta ctatgaatga ataatgctga   101880
gttcataaca tcttttgtaga tatatccaca attttccctc aggataagtg cctacaagtg   101940
gaattactgg actgaaaata atgcagtttg ctaagacttt gctatctgtt cctgaatgct   102000
cctccaaaaa ggttttgcca gtttacatcc tcatgaccag cgaatgagag tgttgcctat   102060
tttcctgctgc ccttgttact gcttaataat ttttgaaaaa aatctaattt gacagacaaa   102120
aatgcatttt atgttaattt gcttttctgg gattttaat gaggttgagt atagtttta   102180
atatttttat tggccccttt ggaactagta tcataagttt ttttttcttaa gaaattatgt   102240
agtctgggct gggcgcagtg gctcacgcct gcaatcccag cactttggga ggccgaggtg   102300
ggtggattgc cgaaggtcag gagtttgaga ccatcctgac caacatggtg aaaccgaact   102360
tctactaaaa gtacaaaaac tagctcagcg tggtggcggg tgcctgtaat cccagctact   102420
taggaggctg agtcaagaga atcgcttgaa cccgggaggt ggaggttggt tgcattgagc   102480
cgagatcgcg ccattgctct ccagcctagg caacaagagt gaaaagtctc aaaaaaaaaa   102540
aaaaaaaaaa aaaaaaagaat ttacatggtc tgaattgcca ttaaaagaaa tatgagaatt   102600
attgagtaac aaataacttt ttaataattt aggcaagttt tggacgattg tacttttgttt   102660
agaaaccaaa agcatagtat ttgtagtttt tttatttact ttagttgcta ggaagtaaac   102720
tttattcaag gtctctggta ccagttgttg ctaaaagtga ttgactaatc tgtcaatctg   102780
aaatttattg ttgctgaact gctaattctt ttgcttctat cttttaggca gacatcttat   102840
ggactaccag actcaagaga ccaaatcaag cctttctaag acccttgaac aagtcttgca   102900
cgacactatt gtcctcccttt acttcattca attcatggaa cttcggcgaa tggaaagctt   102960
ggtgaaattt tggttagagg ctgaaagttt tcattcaaca acttggtcgc gaataagagc   103020
acacagtcta aacacagtga agcagagctc actggctgag cctgtctctc catctaaaaa   103080
gcatgaaact acagcgtctt ttttaactga tttctttgat aagagattgg aggattctgg   103140
```

80

```
ctcagcacag ttgtttatga ctcattcaga aggaattgac ctgaataata gaactaacag   103200
cactcagaat cacttgctgc tttcccagga atgtgacagt gcccattctc tccgtcttga   103260
aatggccaga gcaggaactc accaagtttc catggaaacc caagaatctt cctctacact   103320
tacagtagcc agtagaaata gtcccgcttc tccactaaaa gaattgtcag gaaaactaat   103380
gaaaagtgag tatgtgattt tcttgtgtgt acatatgtgt ctcactttct tttttaatt    103440
tactaagcag aacttcagat gaggaataaa atgattggaa tatttttttt ctcctctaac   103500
tacttgtaaa tttgggagaa tttggagagt gtagtagagt cagatcagtg tatggaaaag   103560
gagcaggagt gactggacct tctaagaagt gtgttatcag aattagtaaa tgaagggtca   103620
aatgtcctac ttttcccctc cactgatttt gacatcaaac cattatccac atagccttat   103680
ttcctccctc ggtcttaatt ttattaatat ttattaagt tcagtgctta gtgtatattt    103740
aaaattttta aaaattgcca ataagtgaca tttattaagt tcagtgctta gtgtatattt   103800
ggattttatt tattagtcac aagacctttg tgcagtgagt aggcatgatt atcttttttt   103860
ttttgagatg gagtcttgct ctgtcgccca ggctggagtg caatggcgcg gtctcggctc   103920
actgcaacct ccgggttcac ggcattctcc tgcctcagcc tcccaaatag ctgggactac   103980
aggcgcctgc caccacaccc ggctaattt tttgtatttt tagtagagac ggggtttcac    104040
catgttcgcc aggatggtct cgatctcctg acttgtgtat ccgcctgcct cggcctccca   104100
aagtgctggg attacaggca tgagccaccg cgcccggact gattatctta tttacacatg   104160
agaaaaccag ggcttagaaa ggctaggtaa gcacctgatt ttcctctag gttgtacagt    104220
ctagaagcat tttgacaaga gcacctgttt tttttttctc tctattagtt tagaaattat   104280
atactcttaa ttatcacctg ggatttttgat tagacagcct tcatgttctt tttcatctta   104340
aatgttcttt gtgtcttaaa gggctaagtg atttcttcag atcttttagt tcactcattc   104400
tcagtgaact aaaatgaggt ctaatctgct actgaatcaa gttttcagca tgttatttcc   104460
ttcctcccctc cctccctcct tcctccctcc aaccaggctc ccgaggagct gggctaaggt   104520
gcgcccgcca ccactcctgg ctaatttta tattttagta gagacggggt ttcaccatgt     104580
tggtcaggct gatcttgaac tcctgacctc aagtgaccca cctgcctcgg cctcccaaag   104640
tgctgggatt acaggcatga atcaccacac ctgacggcat gttattttca tcgcaaagtt    104700
actgtaagct gggagaagtg gcacacactt gtactcccag ctactcagga agcttaaggt   104760
gagaagattg cttgagccca ggagtttga gaccaacctg ggcaacacag caagacccca    104820
gctcaaacaa agaaaaaaag ttattgaatt ttttatttct atggatcatt ttttgtagtt   104880
tcttattcct ttcaccctc attcccactt ttgatcccat cttttattta tttagttta     104940
ttaaatgtat atttgtctga taattctgct atctacagtt ttttgtggac ctgactcagc   105000
atttctttgt ttcttcggat tcagactgtt ggtggcttgt gattttagtg attttttggcc   105060
gtgaacatgt ttcttggact tttgtctgtg ggaattctct gtgtactctg tataaattat    105120
gttacttcag gtgtttttgca ttttcttttg ccatgcacct ggggcctggg tcactaccct   105180
tctggtacca cttaaaactg aattttttgtc ttgggtgctc gtactgatcc tgtatgagta   105240
caggtttata cttactgtag aaatatggtg tttgattatg gggtattgtc ccagatggtg   105300
ctggagtatt aatatgctct ctgttaaact taatgtgttg tccctgtaaa actccaaaat   105360
tctgaattcc agaatactac tggccccaaa tgtttaagat aagggcactg cctgtatttg    105420
tttctgcctc ccactatttt ccttagttta acacaaactc accttttaa aaacattttt     105480
gagagaaattc agtattggga agagtttcta acctgtttct ggaaatggaa gtccaaagtc   105540
tgtttctgta attgttttt ttttgagatg gagtctcact ctgtcaccca ggctggagtg   105600
caatgacgta ctctcagctc actgcaacct ccactcccg ggttcaagcg attctcttgc    105660
ctcagcctcc tgagtagctg ggattacagg tgcccaccac catgcctggc tgatttttgt   105720
attttagaa gagatggggt ttcgccatgt tggccaggct ggtcttgaac tcctgacttt    105780
gtgatctgcc cacctcagcc tcccaaagtg ctaggattat tccaataaa atagaaatag   105840
tttattgttt ttagaaactg tctttgcttt agtggtaatt ttcaataaaa atagaaatag   105900
cagtggagtt attaaaagag cattagttac attttttcct atttttccct ttttcattat   105960
tatatatagt aagttgacc ttttaaaat taaagaacag ttagttcctt caccttttgaa    106020
agattcctgt aactgtcacc actataaggg taaagacact ttagttcctt caccttttga    106080
gtcaagcccc acctctatcc caacacttgg caaccgctga tctttctccg tctcaatagc   106140
tttgcctttt ctcttttt tcttattt ttttttttgag acagcgtctt gctctgtcgc      106200
ccgagctgga gtgcagtgag gcaatctgg ctcactgcaa cctccgcctc ctgggttcaa    106260
gcagttcctcc tgcctttagcc tcccagtag ctgggattat aggcacgcac caccacaccc   106320
caaactcttg acctcaagtg atccacctgc ctcggcctcc caaagtgctg ggattacagg   106380
cgtgagccac tgtgcccaat caggactttt tttttttaaa tttcattca acttgtcatt   106440
tttttctttgt atggattgtg ccttcagagt cacactcaag agccctttgc ctaagcaaag   106500
gtcatgaaga ttttctcata tgtttcctt taaaagtatt gtggttggcc aggtgccatg    106560
gcttatgcct gtaatctcag cactttgaga agctgaggtg ggcagattac gaggtcagga   106680
gatcgagacc atcctggcta atgcggtgaa acccatctc tactaaaaat acaaaaaaa     106740
aaaaaaatta gccgggcgtg gtggcgggca cctgtagtcc cagctacttg agaggttgag   106800
gcaggagaat agtgtgaacc cgggaggtgg agcttgcagt gagccgagat cgcgccactg    106860
cactccagcc tgggcaacaa agtgagactc catctcaaaa aaaaaaaaa agtattatgg    106920
ttttacactt tacgtttaga tatatatctt ttttgagtta atgtcgtata atgatgaggg    106980
ttacgtcaga tttttgttt tttgtttatt tttacatatg gatgtctagt tgttctaata   107040
ccatttgttg aaaagacaac ctttactcca ttgaattgcc tttgtacttt tgccatattt    107100
gtctaggcct gtttttggac tcctttttct gtttcatgat gtgtgtgtct attccttgt    107160
```

81

```
taataccaca  tggtcttaat  tactgtatag  taagtcttaa  aattgggtaa  tgctggcctt   107220
ataaaacgaa  ttgggaagtt  tttatttta  ctcttatttc  cattttctag  aagagattgt   107280
gtagaattgg  tgtcatttct  tctttagata  tttggttgaa  ttgggaagtg  atgccatctg   107340
ggcctagggt  tttgtttttt  gtgtgtgaga  cagagtctca  cttctgtcac  ccaggttgga   107400
gtgcagtggt  gagatcttgg  cttactgcaa  cctctgcctc  ccaggttcaa  gttatcctcc   107460
tgcctcagcc  tcccaaatag  ctgggattac  aagcgtgtgc  caccatgccc  gactaatttt   107520
tgtatttta   atgcagacag  ggtttcacca  tgttagccaa  gctggtctcg  aacttgtgac   107580
ctcaagtgat  tagcccacct  tggcctccca  aagtgttagg  attatagatg  tgagccaccg   107640
tgcctggcag  gggcctaggg  ttttctttt   cagagtattt  taaactatga  attcagatta   107700
tttaatagat  ataggactat  ttaagttatc  tgtttcttct  tgagtgaatt  tttactgtag   107760
tttatggcct  ttgagtaatt  aattgtattg  aattgtcaaa  tttatgagcg  tgtaattatt   107820
tatagcattt  cgggtttgta  gtggtatccc  tcttttatc   ctggtgttgg  caattgtgtc   107880
ttgtttttct  ttgtcagatt  gtataggat   ttattagtct  tttcaaagaa  ctagcttttg   107940
ttttgatttt  tctgttgttt  tgttttcaat  tttattgatt  ttctgctctt  tattatttct   108000
tttctattat  ttctgcttgc  tttgggttta  tttactctt   tttttttct   ccaagttgct   108060
taaagtagaa  acttagattt  ctggtttgag  acctttcttt  tctaagataa  gcatttaata   108120
ctgtaaattt  ccttctaacc  actgctttag  ttacacccc   acaaattctg  gtattttgaa   108180
ctgagcacaa  atgaaaatgtt  ctaatttccc  ttgaatctta  ttcttttacc  aaagattat   108240
ttagaaaatat  gttatttagt  ttgcaagcaa  ttggagactt  ttttcctgtt  attttctac   108300
catttatttc  tcatttcatt  atattatggt  cagagaaatat  attttgaatg  atttcattta   108360
ttaattttta  aaaataacat  taaaaaatgt  ttaaaatgt   gaatatacca  catacagtat   108420
aaagattgta  cattctgttt  ttggacagtt  ttctataaat  gtcaagttga  tttagttggt   108480
taatgatggt  gttcagtttt  tcttttattct  tgctgatact  ttgtatgcag  ttatacacgt   108540
ttattactca  gaagagtgtt  gaacttccca  actacaattt  tttttttccaa  ttttactttc   108600
agctctatct  ggttttgctt  catgtatttt  gaggctctgt  tgttaggtgt  gtacacattc   108660
aggatgatat  cttctgggtg  aattgcctgt  tttatcatta  tgtaattccc  tctttatggt   108720
aattttcctt  gttctaagat  cagaaatatc  tgttgtccaa  ttatataga   cactgcagct   108780
ttcatttgat  tagtgcttgc  atggcatatc  tttttccatt  tttttacttt  tgatctacct   108840
ttataattct  atttaaaggg  ggcttcttgt  aggcagcata  tagttgggta  ggttgtattta   108900
tttatttatt  tattttatta  tttatttatt  tattgagaca  gagtttttgct  cttgttgccc   108960
aagctggagt  gcagtggtgc  aatcctggct  taccacaacc  tccacctcct  gggttgcagt   109020
gattctcctg  cctcagcctc  ccaagtagct  gggattacag  gcacgcgcac  catgcttggc   109080
tgatttttg   tattttttagt  agaaacggat  tttcaccatg  ttagccaggc  tggtcttgaa   109140
ctcctgacct  caggtgatcc  acctgctttg  gcctcccaaa  gtgctgggat  tacaggcgtg   109200
agccactgca  cccggctgag  tcatgttatt  tttaatcttt  tctcacaata  caggttggtg   109260
gttggtaaat  ttaattattt  taatataaat  tttagtataa  ttatttacat  taaatgtaac   109320
tgttgcactg  gggtatttat  aatgtgtaaa  tataattatt  ggtattaata  taattatatt   109380
actcataata  atattaatat  ctttggattt  agattaccag  tttagtatat  gtttttctgt   109440
ttctccctct  ttgatttccc  cttttttgct  tttttttttt  ttttaattct  taattttttt   109500
tagtatttgt  tgatcattct  tgggtgtttc  ttggagaggg  ggattttggca  gggtcatagg   109560
acaatagttg  agggaaggtc  agcagataaa  catgtgaaca  aggtctctgg  tttttcctaga   109620
cagaggaccc  tgcggccttc  tgcagtgttt  gtgtccctgg  gtacttgaga  ttaggggagtg   109680
gtgatgactc  ttaacgagca  tgctgccttc  aagcatctgt  ttaacaaagc  acatcttgca   109740
ccacccttaa  tccatttaac  cctgagtggt  aatagcacat  gtttcagaga  gcaggggggtt   109800
ggggggtaagg  ttatagatta  acagcatccc  aaggcagaag  aattttttctt  agtacagaac   109860
aaaatggagt  ctcccatgtc  tacttcttc   tacacagaca  cagtaacaat  ctgatctctc   109920
ttctttcc   ccacatttcc  ccctttcta   ttcgacaaaa  ctgccatcgt  catcatgccc   109980
cgttctcaat  gagctgttgg  gtacacctcc  cagacggggt  ggcagctggg  cagaggggct   110040
cctcacttcc  cagatggggc  agccgggcag  aggcgccccc  cacctcccag  acggggcagt   110100
ggcggggcgg  aggcgccccc  cacctccctc  ccggatggg   cggctggccg  cgggggggct   110160
gacccccac   ctccctcccg  gacgggggcgg  ctggccgggc  ggggggctgac  ccccacctc   110220
cctcccagat  ggggcggctg  gccgggcggg  ggctgccccc  cacctccctc  ccgacggggg   110280
cggctgccgg  gctgaggggc  tcctcacttc  gcagaccggg  cggctgccgg  gcggagggggc   110340
tcctcacttc  tcagacgggg  cggccgggca  gagacgctcc  tcacctccca  gaggcgctcc   110400
cggtcgggca  gagacactcc  tcagttccca  gacggggtcg  cggccgggca  gaggcgctcc   110460
tcccatccca  gacggggcgg  cggggcagag  gtggtcccca  catctcagac  gatggggctg   110520
cgggcagaga  cactcctcac  ttcctagacg  ggatggcagc  cgggaagagg  tgctcctcac   110580
ttcccagacg  gggcggcccg  tcagagggc   tcctcacatc  ccagacgatg  ggcggctagg   110640
cagagacgct  cctcacttcc  cggacggggt  ggcggccggg  cagaggctgc  aatctcggca   110700
ctttgggagg  ccaaggcagg  cggctgggaa  gtggaggttg  tagggagctg  agatcacgcc   110760
actgcactcc  agcctgggca  acattgagca  ttgagtgagc  gagactccgt  ctgcaatcct   110820
ggcactcagg  gaggccgagg  caggcagatc  actcgcggtc  aggagctggt  gaccagcccg   110880
gccaacacag  cgaaacccg   tctccaccaa  aaaatgcaaa  aaccagtcag  gtgtggcggc   110940
gtgcgcctgc  aatcccgagc  actctgcagg  ctgaggcagg  agaatcaggc  aggaggttg   111000
cagtgagccg  agatggcggc  agtacagtcc  agcctcggct  ttcacaactt  tggtggcatc   111060
agagggagac  cggggagagg  gagaggggaa  cgagggagag  cccctttttt  gctttctttt   111120
ggattatttg  aattttttcct  taaatttatt  tatcttactt  atttatttat  tttttgagt   111180
```

82

```
gattctcctg ccacagctcc caagtagctg ggactgcagg catgtgccac tacacccagc   111240
taattttttt gtattttag tagagacagg gtttcaccat attggccagg ctggtcttga   111300
actcttgacc tcaagtgatc cacctgcctc ggcctcccaa agtgctggga ttacaggcgt   111360
gagccaccat gccctgcctt tttctagaat ttatatattg agttcttgat tgtatctttt   111420
tatgtaggct ttttagtggc ttctctagga attacaatat acatactttt cacagtgtac   111480
tcacattgaa tattttgtaa cttcaagtgg aatgtagaaa acttaaccac cataaaaata   111540
gaactaggga tgaggttaaa aaagagagag aaaagaaatg taataaagat ttaataacac   111600
cgttttttt tttttttctc tttttttttt gagacagagt ctctctttct gttaccaggc   111660
tggagtgcag tggcgtgatc ttggctcact gcaacctccg cctcctgggt tcaagtgttt   111720
ctcctgcctc agcctactga gtagctggga ttacaggtgc gcgccaccat gcccagctaa   111780
tttttgtatt tttagtagag acggtttcac tgtgttggcc aggatggtct cgatttcttg   111840
accttgtgat tcgctctcct cagcctccca aagtgctggg attacaggcg tgagccaccg   111900
cgcccggcta agtctttaaa tatttttttg acattgcact tttctctttt tccttctagg   111960
attttagtaa cccaaatgtt agttttgtta ttgtttgtca ggttcctgag gctttcctta   112020
cttctttaaa tttttttttc ctgttgttca gcttcgaaaa tttctattca tctgtcttca   112080
aattcactgg ttctttcccg ttatttccat tctgttattg agtctttgta gtgaatttta   112140
aattttgttt attatgtttt ttagttctaa aatttttcttt tttgtgtat gtcttatact   112200
ttgctcctga aactcttatt tgtttcagga gtgatcttat ttcttagagc atggttttag   112260
tagctactta aaatttgttt tatcatccca gcatatgtgt cctcttgatt gtcttttctc   112320
ttgtgagata atgggatttt ctggttcttt atatgacaat taattttgga ttgtatcttg   112380
gacagtttga cttacgttac atgattctga atctttgtta aatcctgtgg aaaatattga   112440
agtttttgtt ttaacaagca gttgacctag ttaggttcag tcacaaaatt ctaagcagca   112500
ttctgtcggc tctggttcca tcatcagttc agttttgtat cttatctgct tatgtgcctt   112560
tctgtgtcca gtctgggacc tggccaatgg tcaggtccca aagcctttgt acacttttag   112620
aagcagggcc atgcacacct agctcacgag tggcccaggg agtgcacata caatctgacg   112680
ttttcatggg ctccttcttt tctgtgatgt ccctgacacg ttctgccttc taagaacctc   112740
cctttatccc tttcctgttg tctggctaga aagtcagggc tttagattcc ctatacttca   112800
gcacacttcc tgtagctatg tcaacctctg tggccacgac ttcttcttct tgggactgca   112860
gtttctcttg tcagaaagta ggattcttgg agctgctgtc attgctgctg tggctgctct   112920
gatgctgcct gggagtcgaa ggagagaaag gaacaaaaca aaacaaccca ggggatttcc   112980
tccactctct ttgatccgtg agagccccct ttcctgttcc tcagaccaga aatagagggc   113040
ctgtcttgga acttcttctt tgtgcatctg gtgtgcagtt tcagcttttg agtccaggcc   113100
aggaggtgct ggacaaactt gtcaggagta cggaggtact gcaagttctg attactttc   113160
tcagtccacc tgcttccaag tccttggatg catttgtcca ttgttttgag ttgcattcca   113220
tgggagagac agaagagtgt gcttatttca tcttgacata cttattagga tttcatatca   113280
aatcaacgga tgatattctc tatattaatt tgctgtttc cctttagcaa gcacattagg   113340
aaaataacac tttaacaccc gcctttggtg gtttctgtca taattattaa tactttgactt   113400
tttttttttt tttgagacgg agtctcactc tgtcctttga ggcattgtcc ccataaactt   113460
ttggtaaagc atcaataatt ttatctttca tccacacaag cttcaccata aatttgatgt   113520
ttattcttcc attttagcag aattcatgtt gctccaatag gggctgtctt caaactgatg   113580
tttttctcctt cttagtgcct cagagtagat cctgttcaga tacgttataa caggttaata   113640
tgagtttatt ttggtgtaaa agtactttga aattcatgca tagtttttttc atcatatgca   113700
tttttccatag ctttgaacac ccccatgtaa ctctcctctt ccacaaacca aacaatgaaa   113760
aagcaccttt gtgatcggaag tttattttgc aatagggaact cacagtgatc taagccctgc   113820
tattcatgaa tataatttcat tactggagtc caagttgctt tttggttttt gaagttctct   113880
tcttccctta caggtataga acaagatgca gtgaatactt ttaccaaata tatatctcca   113940
gatgctgcta aaccaatacc aattacagaa gcaatgagaa atgacatcat aggtaagcag   114000
tgcttgaaac tatggcaaaa aaaaaatgac aaaaaatgca cagaactgac aatttttcgtt   114060
attgactaag ataattttt cttaacatgg aatttagcag ttcccttcct aatttgtttt   114120
ctgagtattt tttatatcgg attatagctc acttttaaaag tttctcggct ggcttcggtg   114180
cgagggtgtt tgcctgggcc agatgggctg cagtgtagcg ggtgctcagg cctgcccgct   114240
gctgagcagc gggagccggcg ggcggctacg ctaaccgaccg cagaccaccg gatggactgg   114300
ccggcagccc cgcaccagtg cacgaagtgg gcgggacaga aacttctggg gttggaagtc   114360
cagtgaggct aaaagccggt accaaagtct ctaggcatca gggctgcagc ccaagagtct   114420
cacgaccagt gggcaactgg atggccagac aggtgtctca gtggtggcct ctccgtctca   114480
gggcttcatc ccacttctca gtgggcctga cgtccctgga gtctacctgc   114540
attagccaga gccatcacat ggcctgtgac ttgccttttt ttgccagttg attgtgccac   114600
acacagtgtc atttctgtgt catttggcac agctggaggt gcaaggagga gggcgcagcc   114660
atgtccagtc ccagtttcac gtaacttttat tcttctgaat aaagacaatt tgctaacctt   114720
aaaaaaaaaa aaaaaaaaaa agttttttctt atatgttgga cccaaatctt taggctttaa   114780
cctgaataac aatgacagca agatcaataa atagtacaca tttattaaac actcactgtg   114840
tcccagacaa tattccaagc acttttatg gatagactca ttttaacttc taaagaactt   114900
tgtgggataa atacagttat tttatatgatg aagaaactga agcacagaga agttaagtgc   114960
tttgtccgag gtaacagctc agatatggca gagtcaggat ttgaaactag acccctcacat   115020
accttaactg ctgtgctgtg gcagtgtttt tcatactgta ggttgggacc aatgtgtctc   115080
tatgccctca cccctgcca aaaaaaaaaa aaaaaaaaaa aaatatatat atatatatat   115140
atatatatat atatatatat aatatatata tatataaaat atatatat ataaaatata   115200
```

83

```
tgtattagta tatatgcata tatagtatat attatatatt agtatatata ctaatatata  115260
atatacatat tagtgtgtgt atatatatat atactagaat aaaaaaatca aagtatctca  115320
gagtagtaag gacaaacatt tcagaaaaat gttttcatta tatatacatg tatgtatgtg  115380
tatgctgatt caacaaaaat atttcttata ggttatagca aaatagtttg aaagctttta  115440
ctgtgtttta tcaggaagac cttaggtgaa cgtatattca cagatataaag aggttattta  115500
ttcattcaat aaatattaca ttctcataag tcctaatatt atgtatttt atttttcaaa  115560
aaagttagta tttgtgattt atgaaataag acatgttctt gcactttag cagatctgt  115620
ccgatgttgg gcttctttaa tccttagtgt gggtgctttta cactcactca ctgctgggga  115680
cagcaagacc cctgttagtc tcagctgtgt ttcttaaatt ggcccactgt accttccagt  115740
tagctattct ggggtccatg tcatgttgag tccatttcc ttttctttct cccacacaga  115800
tacctataac ggctataaca taggcctggt ggctgtggt ggcttatccc tatctgcttg  115860
tatttaaggg gtactgtttc actgagtttt gctgacagat gttgtcatga gatttgaggt  115920
tttctgtgtt gttgctctat ttttatgtgg gaattgcta ctatcatcat ccctacgacca  115980
gcttttccta gtaatacaac agggatgttc tgactgatta gagtttgcct gtttgaagaa  116040
ttggttggct agtgatttt tttgaggag agtctgtacc agttaatgac ctgactggcg  116100
tgtggataaa aaggaagcag tttcaagtca aataaaacac ttaaaatgaa accacactgc  116160
aactctcttt cttttactta agcttaatca aattaatgat gatgtaatcc catgaaggaa  116220
aagtcttctg aaggatcaag ttgataacat tttgtgatca aagaatttga gaaaacctct  116280
atcccagtgt ctatcattat atatttagg atgttaatta cctgtgtggc tttaggcaag  116340
tcatttttct tccttgagcc ccattcttaa tcctgtccaa attatttgtc tcctcttgca  116400
gttgactat tttaatacag ctgtccttca agtgagtttt gttcaaagga acctatact  116460
tagctcttac tgtgtaccca ctttgcatag tcttgtttta aatgtaatcc ttggatttt  116520
ggtgttgcta actaattact gtttttatgt gaggatttag agtgatccag aatctataact  116580
tgcactacct ccttcatctt ccacaaatgt ttgaagtggt agaattttta aaaactttga  116640
aggtacagct gacagaattc gctgatggtt tggaagtgag tggtatgaga gggaaaaaaa  116700
ggaataaagc atgactgcat ttttttgtttg tttgtttgtt tgttttttgag acggagtctc  116760
actctcgcca ggctggagtg cagtggcgtg atcttggctc acggcaacct ccgcctcctg  116820
ggttcaagcg attccctgc ctcagcctcc caagtagctg ggactacagg cgctcgccac  116880
cacgcctggc taatttttt ttttgtattt tagtagaaac ggggtttcac cgtgttggcc  116940
aggatggtct ccatctcctg acctcatgat ctactcacct tggcctccca aagtgctgag  117000
gttacagcca tatataaag catataaagt gtgttatagc atacaaacag gtatatatat  117060
aaacatgcag tccacacagc tgatatggaat gaggcagtag tgaaggagaa gttgatgtag  117120
gagagggggac agttgttaca ggaaagaagt ctggaggcag aagggatgaa ttccagtgct  117180
cacatagaag attgcttaga tgggagcaag gacaatttat ctagagtcac aggaaagaat  117240
gcagtacacg ggtagaagatg caggtgagtt gaaagatgtg agagatgatg gaaataattt  117300
tctgattgct tctatattct caaggaagca ggaagcaaag tcctcagcaa agagaataga  117360
agaggtgtta aatatttgag aaaggagatg tactgtagaa aaaaaaaaaa ctcatttctt  117420
ccttctgaac tctcacaaaa cagaacccct ccatgactct agttgtgtgg ggtttttcc  117480
ctgtcagcta ccaattctgc agatgattgt tcagtgaaca ccaactgggt gtcctctaag  117540
tcagttcagt tctcacactg tttacctgga gatagcatca gatcccacag attgaggact  117600
ctgtcccaca agactgcctc cacttcagat gccagtctca agtacaagtt ggggcctgtg  117660
cttctgactg accttctata aattggagtt cccacagtcc cctccttggg ttcaataaat  117720
ttgctagag agctctcaga actcagggaa atgctttaca tatatttacc catttattat  117780
aaaggatatt acaaaggata cagattgaac aggcagatgg aagagatgca tgggcaaggt  117840
atgggaaggg ggcacaggc ttcatgcac tctccaggtc atgccaccct ccaagaacct  117900
ctacagattt agctattcag aagccccccct cccatctctg tcctttttggg ttttttgtgg  117960
agacttcatt atataggcat gattgatcat tggctattgg tgatcagctc aacctcagc  118020
cccctcatcc cgggagggttg gtgggtaggg ctgaaagtcc caacatgta attctgcctt  118080
ggtctttctg gtgattagcc ctcatcctaa agctcttag aggccacagc cacaagtcat  118140
ctcattagcc ttcaaaagaa tccagagatt ccatgaattt aggcgctgt atgctaagaa  118200
actggctaaa ggccagttgc aatgtctcag gcctgtaatc ccagcacttt gggaggctga  118260
ggcaggagga tcgttcagg ccatgagatc aaaaccagcc tggtcaacat agtgagaccc  118320
cctttaacaa aatttaaaaa ttggccaggc gtaatagtcc ttgtctgtag tctcagctac  118380
tcagaaggct gaggatcact gagcctggga gttgaaggca gcagtgagcc gtagtgatgc  118440
cactgactcc ggcttgggtg acaaagtgag accttgtctc agaagaaaaa ggaaaaaaaa  118500
aaaactgggc aaagactaaa taacatattt cacagtatca cagatttgta ttgtctagga  118560
aagtgaatgt aaacagacca ggacactagt atgatccctt ggtttcatga aggtcccact  118620
aaagtcatga acacaaagtg agactaggca tcatgttata tggtttttcc agccatgttt  118680
aacagctagc taaatagcta attgtttcgc tgcagtttat tttagcagtt ccttattta  118740
gcacatttca tgttttaaaa tttctaccaa taacatttta ataaactttt ttacagataa  118800
cttcacaaat ccataatttt ttaagttaca atcccagaaa tagaattgct cattgaaagg  118860
gtatgttcat ttttaaagtt atgctagaaa ctgccaaatt gccttcagaa aaaggtgttt  118920
gtatcccac taacactagt gttagtttc ttgtgccctt gctcaagtat acatattatt  118980
aaaaacaatg ttgggccagt ttactagata aaaggtgtag tgcctcctta ttctaatcta  119040
tttgattact agtgagtatg tatgtctttt cacgttggtc attttatgtt tgttcctttg  119100
tggattgtca tgtcctttgc tcatttttct tttggaacat ttcttagtag tttataagag  119160
ctcttggtat tttaatgata gtaacctttt aactgtcatg catgctgcaa atctttttc  119220
```

                                    84

```
tgtttgtttg cctttgtatt ttgtttttgg agggtttcta tgtataggaa ttaaatttta    119280
tgttgttaaa tcttttgatt tctgcttttg catatgtact tcaaaagact ttctattta     119340
agatcaagtg ttacctgtat tttctttag ttctatttaa aacctcttaa tttatatgcc    119400
tgtgctgtta actcccaagt tgattcacaa gtgtgtatac atagtttgaa tttagtggca    119460
atttaattat ttacaacttc ttttgcagca aggatttgtg gagaagatgg acaggtggat    119520
cccaactgtt tcgttttggc acagtccata gtctttagtg caatggacga agaggtaagt    119580
agttcatatt ttcacattgt gcatcctagg gaatttgggt tcattgttag gaatgggctt    119640
cactcagcta aaaacaaagt attttgaga atttaaatat tttgatatt tacaagatca     119700
tataaagcat actctatctt ggttaacagt ttcttttaaa tataaattat gtgaactctt    119760
aaaattttca ttttcatttt caatgttaat atttcctaag ttaaaataat ttgtttttag    119820
ttctgaaata atttggggag tgattgagtc tgtagtgatt atgactatta gaattggttt    119880
atttatttaa ataatgcatg tcttcagatg gctctcctaa tttgttagtt aggcttgaag    119940
ctaaatggat gctatataac taaatccaca tagatttgtt gaaatggctc cagaggtttt    120000
ttagatttat tactgctatg tgcccttaaa aaaaatctat tcattctttc acttaacatt    120060
tatcagaaga gtgctctgtg taagacgtgg ttaggcatag tgccagtctt gaaggaagtt    120120
acagcctaat aaaagacata gggcatgttg tttggttact gtaatatgaa gtggcagtg     120180
ttaaatgtca ggggagaact acaaagtcat aaaaaggtgg gagagattac atacaggtaa    120240
aggaatcagg aatgacacca tggggagtaa ggtagtgttg acctaggcct ttaagatcaa    120300
ataggacagt tatggaaaga gtatatttt cccacttaaa ctctttcctt ggtcgttccc    120360
tcaaattttc cctttgtcc atgtgcaggc actttagtga gtttctgcga agtcaccatt    120420
tctgtaaata ccagattgaa gtgctgacca gtggaactgt ttacctggct gacattctct    120480
tctgtgagtc agccctcttt tatttctctg aggtaaagtc tgcatttctt ttcacactct    120540
attcgagcat tccagcctct aactatcaat gctgggggcc tgtctatagg aaaatcacca    120600
gaagagccaa gtcatttcca aaaagatgta tcattgtttc aagttgtttc tgatggcaag    120660
agtaatttaa taatatatta gagagaacat gaaaattcaa tgtattaaat aactctaatt    120720
ttgagaaacc taattaaact actgcatgta agagagtgca tgtttttaat tatttggagc    120780
tatttaaaa ccacagaatt tgaaacttgc ttccagtgca taaattgcag accagacttc     120840
agaagagaaa aaaagtagta aatttttct tatgctcatc atttttactt tagtcacttg     120900
ataggattgc ccagtgaaga agcatttgca acagacaatg agtatattaa tcttttttag    120960
gcatacagtt tagtataatg ctctttgtta ggcttcaaca agtgaaatta ttttgttgga    121020
aagcaaatga ctattaagta gaaagaggat tcccagtctc acaaagcagt aatttagaca    121080
ctcgattctg cctctttaca agaatacagg tactcagttg atttgtttc tcactccctt     121140
tctttgctat aagtttaaat caacaatttg tttaggttaa tatgtcctca tggaatggtg    121200
gaaatgatca gatataaaat atttggtttg gttagtttac tctttatatg tttgctggca    121260
aggaaccaca aatccagatt agtataattt ttactctagt tcactaaaag tttgtcatcca   121320
gctgtgtagg tagtgtttgt ttcttgttaa ctttttttttc gtctaaaaga atactttaaa  121380
actttttcaat ctcaaatgac tgtaacttgc tgacaggtgt taacagaaga agtagatctt   121440
tttgttttt gcttatgacc tgtatttaa tatttgagct tatagattag agattgttag       121500
agaatctgat ttatagtctt attttccctt gtgtatttt tcttcctagt acatggaaaa     121560
agaggatgca gtgaatatct tacaattctg gttggcagca gataacttcc agtcctcagct    121620
tgctgccaaa aaggggccaat atgatggaca gaaggcacag aatgatgcca tgatttttata  121680
tgacaagtga gttatattga tagatggatt cagcagatac ttattgaaca tttgatatgt    121740
tttgtgggaaa taaagatgaa taaactcagt ctctgttgtc aaaggagctca caggaggcag  121800
cataaaagct gctttttatat ggtgtttgtg aagctttggg ggttcttaga acaaagtt     121860
ctgctgggaaa aggggaggtg tatgtggggt aaacaggatg gcaatgatgg tgttcaagga   121920
gtgtttccca gaagaggat tttgtttgga tcccaaagaa agaagggaat tttgctaccc     121980
agagaaggca gaaaacaaca ttctaggcaa aggcattggc ccagagacca tggaaacgta    122040
ggggaaagtg gcacttttcaa gaaacttgag tttagataat caaaggagtg gggaataaat   122100
atgaggatgc tggtactaat tggaatagat tgtaaggggac cttgaatgcc tatttatggg   122160
tatattatac ttttctgtata aatctgctca ggcacgttgt taattagttt ttattactgtt 122220
ttcactgaaa atgagaggat ggaaacatca tacagtaaac aaaattgaaa atatctggtc    122280
aggcagatga tgagcttgtg gccagcctgt taacgtatgg tattcttttc atttaactt     122340
tcttactctg taaaaaaagt aattcgttgt cgggcacggt ggctcactcc tgtaatcaca    122400
acactttgga aggcagaggc aggtgaatcg cttgagccca ggaatttgag accagcctgg    122460
gcaacatggc aaaaaccogcc tttactaaaa atacaaaaat tagctgagcg tgatggcgtg   122520
cgctgttgt cctagctact tagggggcctg aggcagaagg atcacctgag cctttgggagg   122580
tcgaggctgc agtgagctgt gatccactgt actccaccct gggcagggca gtagagtgag    122640
accctgtctc caaaaaaaaa aaaaacaaca aaggtaattt gttatttgta tccttaagca    122700
aatgctaaag gggtaacttg gggatagaga aaagtccaca gatgttaggg tttgaagaca   122760
ctaatagtat ctaggccagt ggttcctgaa cattagtctg tgggctcttg ctgggctgctc  122820
tgcataggaa tcacctgaga gcttattaaa aataggtttt caggctggtt gcggtggctc   122880
acgcctataa tcccagcact ttggaggcct gaggcaggcg gattacttga ggtcaggcgt    122940
tcaagaccag cctggccaac atggtaaaac cccgtctcta ctaaaaatac aagaattagc    123000
caggcatggt ggcacaccct gtaatcccca gctactcagg aggctgagga agggaggattg   123060
ctcgagcccg ggaggtggag gttgcagtga ggagatca tgccactgca ctccaggcctg     123120
gctgacagag ggagactctg tctcagaaaa aaaaaaaaa ataggtttt gagtctgggta     123180
ccggtggctc acacctgtaa tcccagcact ttgggaggcc aaggcaggca gatcacttga    123240
```

85

```
ggtcaggagt ttgagaactg cctggccaac atagtgaaac cttgtctcta ctagaaacta 123300
caaaaaatta actgggcatt ttgacgggtg cctataatcc cagctactag ggtgctgag  123360
gcaggagaat tgcttgaacc cgggaggcag aggactgcat ctcaaaaaaa aaaaaaaaaa 123420
aaaggtttcc agtcccctg tctcagaaat tctgattctg caggtttgag gtgtgaccag  123480
gaatctttat ttttagaaga cataccagat aattctgata aatagccagt ttaggagtgt 123540
agtctaattt tcctattttg caagtaagga aaataaggcc cagagaggta atgattttct 123600
caaagtcaca gaacaagtta gtgggcagaat ttggactgga atgcagtcct taatgttctg 123660
tccagtgttt attctggtac agtatgtttg agcagttgat tacgtaagaa acattgttat 123720
atagatgttg agataggaag agtttacatt tagaaatttg gtctaaaatg cctgaacatt 123780
caagtcgtgg aggagtattg accaacttac tcaatacaac ataggagatt cacattttgt 123840
tacaaaaatg tctgatttaaa aggagagttt tcttttttt cttctttttt attttttgag 123900
atggagtctt gctctgtcac ccaggctaga gtgcagtgac acgatctcag ctcactgcaa 123960
cctccacctc ctgggttcaa gcggttctcc tgcctcagcc tcctgagtag ctgggattac 124020
aggtgggggc caccacgccc agctaatttt tgtattttta gtagagacag ggtttcacca 124080
tgttggccag gccggtcttg aactcctgac ctcaagtgat ccacccaacca ctgcctccca 124140
aagtgctggg attataggcg tgagccactg cgcccagcct gcttgttttt gtatcatata 124200
tatgcatcat cataatcatg caatatgatg cattgtattt ctgtcaggac atagaaaaca 124260
ttagagtgct tggaagagga ccttttttt ttctctcgcat ttaatgcttt ttttggtatt 124320
catttcataa tcagcttacc aaaacattac ctgcattata cccatcaag gtagaaatct 124380
ttgtgttatc aatattggtt actccctttc cacaccgagt catcagtaag tcctgttcta 124440
ttcaaatagg tcatatgcat ctagctcacc cctcagtgct gttttgtttt gaatttgtac 124500
atgtttactc ctgatgcctt gtagttatga tgatgtgttc ttatttatt ctgtgcatac 124560
aagttctcag ctcgcttttt agggaaaatg accatgtctt ccttttcctat aaattcctttt 124620
ctatctatca agtcctcaac agagaataag tacccataaa tatgtgattg ttagtttctt 124680
tgcctcagtt gtagtctgat ccttacagct tttaaacaac agtagagttc accgtcaaga 124740
actaaggatg gttggcaggc agatagaaag gtagcaagtt gacccaacta tctctgggga 124800
agtgggaaca aagaaaggtt acatcagcac tgtcatcaca tagctctata gttctaggcc 124860
tgcaggctca atcaagtagc cttgtataag attctctgga ggaggtgctg aaagttgctt 124920
atacttgcta tggaatttta tttttacttcg gatatctttt taccataggt acttctccct 124980
ccaagccaca catcctcttg gatttgatga tgttgtacga ttagaaattg aatccaatat 125040
ctgcagggaa ggtgggccac tccccaactg tttcacaact ccattacgtc aggcctggac 125100
aaccatggag aaggtaaccc agaacttcaa acgtatcaaa ctacaagaag ttttattggt 125160
agaactcata aaatataagg tgggaaaacc aagcagaata gcacagtgga aattgaagca 125220
gtccagcaaa gtgattaaga gcagaggcct tgagtctggc ctggtatgta cagtcacgtg 125280
ccacataaca ttttagtcaa cagtggactg cgtgtacgat ggtcctgtac gattataatg 125340
gatcaaagct ggtagtgcaa taataacaaa agttagaaaa aataaatttt aataagtaaa 125400
aaagaaaaaa gaaaaactaaa aaagataaaa gaataaccaa gaacaaaaca aaaaaaatta 125460
taatgagct gaaaaaatctc tgttgcctca tatttactgt actatacttt taatcattat 125520
tttagagtgc tccttctact tactaagaaa acagttaact gtaaaacagc ttcagacagg 125580
tccttcagga ggttttccaga aggaggcatt gttatcaaag gagatgacgg ctccatgcgt 125640
gttactgccc ctgaagacct tccagtggga caagatgtgg aggtgaaaga aagtgttatt 125700
gatgatcctg accctgtgta ggctaggct aatgtgggtg tttgtcttag ttttttaacaa 125760
acaaatttaa aaagaaaaaa aaaattaaaa atagaaaaaa gcttataaaa taaggatata 125820
atgaaaatat ttttgtacag ctgtatatgt ttgtgtttta agctgttatg acaacagagt 125880
caaaaagcta aaaaagtaaa aacagttaaa aagttactag gtgtaagtct acagtagtgt 125940
aaaaaaattt taaatataat tagtgtagcc taagataca gtgtaagtct acagtagtgt 126000
acaataatgt gctaggccct cacattcact taccacttcact tcgctgactc acccagagca 126060
tcttttatca tgtatttta ctgtgcttt tctgtatttg tgtttaaata cacaaattct 126120
taccattgca atagtggcct acgatattca ttataataac atgtgataca ggtttgtagc 126180
ccaaaagcaa taggttgtac catatagcca aggggtgtag taggccatac catctaggtt 126240
attgattgat tgattgattg cacactgcaa ctctgcctc ccaggttcaa ccaattatcc 126300
tgcctcatcc tcccaagtag ctgggattac aggcaggcac caccatacct ggctaatttt 126360
tgtattttag tagagacagg gtttcaccat tttgttcagg ctgttctcga actcctgacc 126420
ttaagtgatc tgcctgcttt ggcctccgaa agtgctggga ttacaggcat gagctaccat 126480
gcctgggcag taactgaaat tctctaatgc catttttct atctgtaaag tgacgataat 126540
atgcacgttt acctcaaagt tactttgatg attaaagtaa ggtaatgtat ataaaataca 126600
tattaacata gtacctgaca catggtaagc atcaaaaat gttaactact tttattacta 126660
ttattattac gtattttaa ataattagag agcagtatca aaaattagct gggcggtagtg 126720
gcatgcacct atagttccag ctactcaggga ggctgaagct ggaggattgc atgagcctgg 126780
gaattaaagg ctgcagtgag ccgtgtttcat gcccctgcac tccagccttg gtgacagagc 126840
aagaccctgt cttgaacaat taaagaaggc attatgcgcc aacgttagct taagaatgat 126900
ccacatatat caccagtaac tgtcaacagg attggaaccc tagtttttggg tattatgatc 127140
acaaggtatt attaatagct tattaataat aaagcgttgg ctaggcacgg cgactcacat 127200
ctgtaatccc agcacttggg gaggccgagg tgggtggatc acctgaggtc aggagtttga 127260
```

86

```
gaccagcctg accaacatgg agaaacccca tctctactaa aaatacaaaa ttagccgggc 127320
gtggtggtgc atgcctgtaa tcccagctac ttaggaggct gaggcaggaa aatctcttga 127380
acccgggagg cagaggttgc agtgagctga gatcgcacca ttgcactcca gcctgggcaa 127440
caagagcaaa actccgtctc aaaaataataa ttataataaa taaataaaag taaataattg 127500
atgtttgtga atgatttatt cttctaatga actagaggag attttttccag gaatttcaga 127560
gccagtgagg ttatgttgct tgtatgtgtc atgtgtatcc aggtgaaaaa acttaattaa 127620
acgctattat ataataccat acataaaaac tgaattttag gaatactgaa gaatgacata 127680
tagaagtcaa atcattaaat agctagtagt aaacagaata gagtgtcagc tgttacccaa 127740
tgatgataat attttcacga ttaaaattaa acctttctg attttaaagg aaaagttcag 127800
atctgtatca tataaagaat gtaaattttc agggtaataa aattaaaatg cagagaataa 127860
aatgcaaaaa tagttcttac tagatgtgtg tatgtaagga acttagacta attttaagaa 127920
cactgtcaag accctggtag ttaggtagga aaaaagacat gaatgattca ttcaacaaaa 127980
actttgagta tttctgtgct agatggtagt gttacagtgg taaacaaaat aaatgtgttt 128040
ctgctatcct ggagcttagt ctacaaaaaa ggtacatatt ggccgggcac ggtggctcac 128100
gcctgtaatc ctagcactt ggaagatcga ggcgggtgga tcacctgagg tcaggagttc 128160
aagaccagct tggccaacat ggcgaaaccc cgtctctact aaaaatacaa aaattaactg 128220
ggtgtggtgg cggacacctg taatcccagc tactcgggag gctgaggcag gagaatcact 128280
tgaacctggg agacagaggt tccagtgagt cgagatcatg ccactgcatt ccagcccggg 128340
ggacaaaagc gaaaatacgt ctcaaaaaaa caaaaacaaa caacaaagc acgtattaaa 128400
tacgaacata aatattatca aattatactg aataagttct catgttatt atttgcttgt 128460
ccagttacaa acttttcctt cgtagaatta gaaatataaa taataacat gagaactcat 128520
tcagtataat taataattat taaatgtaaa taaaaacatc tatgtacaat taggcattta 128580
tttaagaatt atttgaaaaa aaaacaatgt ggaaacagat attttgatat attgctagtg 128640
attgaaatg ataatgttct tttgaagagt aaagtgacca tatatattaa agttaaaatt 128700
taactcagca atcacacgcc tggtgagtta tcttaaggaa atcagtttga aagtaaaatc 128760
aatatatgca caaagacttt aacatttatc ataaaccaga aaaatcgagt ttcaaattat 128820
atcctatgga ctattttctg ctaaaaagta ttaatatcaa ctttatgtaa tactttcgtg 128880
acaaatattt tgggggagaa aacccaacaa aattacatgc attgtaatt ttttttttt 128940
tttttttta gacagtcttg ctccagcgtc caggctggag tgcagtggtg caatctcggc 129000
tcactgcaac ctccatctcc caggttcaag caattctcct gcctcaggcc tcccgagtag 129060
ctgggattac aggcgctcac caccatgcct agctaatttt tatagttttt agtagagatg 129120
gggtttcatc atgttggcca ggctggtctt gaactcctgg tctcaagtga tccgtctgcc 129180
tcggcctcct agagtgctga gattacaggt gtaagccact gcacccagcc ttatgcatta 129240
taattttaat ttgtaaactg tacaaaggga taatacttgt agtacaacaa gaagtaaaaa 129300
catttgttat aggtagttaa catttgtaac cagtagaatt ataggtaaaa tttatttatt 129360
taaaacagtt ttagttggat ttgatttcaa ctttaaaata atgctttttca tctctatcag 129420
gtcttttgtc ctggcttttt gtccagcaat ctttattata aatatttgaa tgatctcatc 129480
cattcggttc gaggagatga atttctgggc gggaacgtgt cgctgactgc tcctggctct 129540
gttggccctc ctgatgagtc tcacccaggg agttctgaca gctctgcgtc tcaggtattg 129600
actgattgcg tctgccatta gggagaaaag catacacatc ctttccttca catcccagta 129660
acagatccta ttatttgtaa atttaagtt gtggaaaaaa aagataaaag ccaggcacag 129720
tggcctgtgc ctgtaatccc agcactttgg gaggctgcgg tgggcggatc acactgaggtc 129780
aggaattcga gaccagcctg gccgacatgg tgaaacccca tctctactaa aaatacaaaa 129840
attagccggg catggtggca ggcacctgta atcctagcta cttgggaggc tgagtgcagga 129900
gaatcgcttg aacccaggag gcagaggttg caatgaacca aaatcacgcc actgcactcc 129960
agcctgggtg acaaagtgag actgtgtctc aaaaaaaaaa aaaaaagaga gaaataaaat 130020
tagcctactt actatcttct aatcaaagca tttgtggtaa cttaaaatat actgtattgt 130080
aaagtatcat gctgtttcat ttaggccatt attctatttg aatctgtggc tgttttctctt 130140
aataaatcaa gtaatatgga atatattcat agcctctgaa gagctctta tgtaagtatt 130200
tatttaggat acttttttgta aaataagtga atgaattctt aggtctcctt tttttttctt 130260
ttcttgagac agggtctcct cgctgcaacc tggaaattct gggctcaaat aatccaccca 130320
ccacagcctc ctgaatagct gggactagag gcatgcacca ccacgcctgg ctaatttgaa 130380
atttttttttt ggccaggcat gatggttcac gcctgtaatc ccagcacttt gggagacccga 130440
ggcaggcaga tcacgaggtc gggagatgga gaccagcctg gccaactgg cagtgagcccg 130500
tctctactaa aaatacaaaa attagctggt tatggtggct catgcctgta atcccagcta 130560
cttgggaggc tgaggcagga gaatggcttc aaccagggag tcggaggttg cagtgagcccg 130620
agatcacgcc actgcactcc tgcatggtga cagagtgaga ctccatctca aaaaaaattt 130680
tttttttaaa tgatggagtc ttgctgtgtt gctcaggctg gtcttgaacc cctgacctca 130740
aatgccgcct gcttcagcct aagtttcttt ttttttttgta aagagacagg gtcttgctat 130800
gttggccagg gtagtctcaa actcctggct tcaagcagtc ctcccacctt ggcctctcaa 130860
agtgctggga ttacaggcgt gaaccactac ctataatgtt gtgtttcact caaggccttt 130920
tgatttcgtt ttgcattacc gtgccacatt gtgcatttcc ttgaccttt ttggggtttt 130980
tggagtgctin tcatatgtta aaccataccct gattctcctc aaaatcacac aaagtagaat 131040
atcctaagac aagaaatcta aggaggcata aagaagttaa ctggtttat taaactcaca 131100
cagtaaatga tagagccaga aatattcccc ttctagtgtt cttcaccatc agcttaatgt 131160
agcataataa tttttctaatt actgttgaca aataaataac cctttgaatt ttcaatactg 131220
ggccttggat aaattttcct aatttgtaag agagtattat cgtattgcca ttacaaaagc 131280
```

87

```
tctcctgagt atcttttct tctgttaagt ttacctagga gataaactgc tgagtatggt    131340
tgccattttg gtttttgat ataggttaga atgtcttggt tttttttttt ttttttttg    131400
gttttttgttg ttgtcattgt ttgagacagc atcttgctct gtcgcccagg ctggagtgca   131460
atggcacgat cgtggcctac tgcaacctcc acctcccggg ttcaagcaat tctcctgcct    131520
cagcttcctg agtagctggg attacaggca tgtgcaacca cacctggcta attttttgtgt   131580
ttttagtaga gaagggggttt caccatgttg gtcaggctgg tattgaactg ctgacctcat    131640
gatccacctg cctcggcctc ccaaagtgct gggattgcag gcatgagcca ctgcacctgg    131700
ctgaatgtct tgttttttgat taggcactta agaaaggcct aggtactaac cataaaatat    131760
atttttatac cttttgttga tactatatat atagaaaact gcacttatca taaccttaga    131820
cacttgaag aatgttcaca agcagaacta acccatgtca cccagcatcc agatcaaaaa    131880
cagcattatc agcccctcta gaagccctct tgggcccctt ccattcactg tccttcttgt    131940
caccagggta gctactatcc tgacttttga tggcatagat tagcattacc tgttcttgtc   132000
attttataaa taaaaccata ctgtgtattc ttttcttgta cagcttttatt gtgctaattc    132060
acatttacat catacaattc agtggttttt atatggtcac agagttaggt aaccattacc    132120
acatcgattt tagaacattt ttttcactcc agatagaaac ccccttttact taaactccaa    132180
atcccccact ccaccagccc taggcagcca ctagtctact tttttatctct atagacacaa    132240
tagattttgct tattctggac atttcataaa catggaacctg tatattatgt ggtctttttgt   132300
tgccaactgt ctttcacttta gcatcatgtg ttcaaaagag catcatgtta tccatgtttg     132360
gcatgtatca gaatttttatt cctcattatg gccaaaatatc ccattgcaag gatttatgac    132420
atttatttg aattgtaccc tcctttctgc catttatcaa taatgctact ggtgaccattt   132480
gtgtacaagt ttttgtgtgg atacaggttt tcttttttgt tttaaatttg aggtgggagtc    132540
ttgctctgtc gcccaggctg gagtgcagtg gcacaatctc ggctcactgc aacctctgtc   132600
tcctgggtcc aagcagttct cctgcctcag cctcccgaagt atctgggact ataggcaccc    132660
accaccacgc ccagctaatt ttttagtaga gatgggggttt caccatgttg gccagtctgg    132720
tctcgaactc ttgacctcaa gtgatccacc catctcggcc tcccaaagtg tcgggattac     132780
aggggtggac cactatgccc ggctgtggtt ttcatttctt ttgttgtata tacataggag     132840
tagaattgct gagtcaagag gtaactctta aacttattga aaaactgcca gattgttttc     132900
cgaaaaggct gcaccatttt gcaatcccac cagcagtgta tgagtttttac agcttctcca    132960
catttcattg gaacttatta tctgtttggc tgtttttaaa aatgatagtc attcaataaa      133020
gttctacttc agtgtggttt ttgcacttct ctgatgagta atgatgttga gcatctttttc     133080
atttgcttat tggcctttgt tctagctttg gaaaaatgtt tattcaaatc ctttggccat     133140
tttattttt attttttatttt attttattttt ttttttgagac aagtctcact ctgtcagcca    133200
ggctggagta caatggtgtg gtctcagctc actgcaacct ccgcctcctg tgttcaagtg     133260
attctcctgc ctcagcctcc cgagtagctg ggattacatt tcaggcacct gccagcatgc    133320
cgggctgatt tttgtattttt tactagtgac agggtttcac catgttagcc aggctggtca    133380
caaactcctg acctcaggtg atctgcctgc ctaggcttcc caaagtgctg ggattacagg    133440
cgtgagccat tgggcccagc ctagattttc ttttttcttt tttttttga aaggggagtct    133500
tgctcttgtt gcccaggctg gagtgcaatg gcacaatctt ggctcactgc aacctctgcc    133560
tcctgggttc aagcgattt cctgcctcag cctccccaagt agctggggatt acaggt atta    133620
accaccacac ccagctaact tttgtatttt ttttagagac agggttttac catgttggcc    133680
aggctggtct caactcctga cctcaggtga tccacctgcc ttggcctccc gaagtgctgg    133740
gattaccggc atgagccac aggcccagcc aatttctca ttatattgcc caggctggtc     133800
tcaaactcct gggtcaagt gatctcctg ccttggcctc ccaaagtgtg gggagtacag    133860
gcgtgagcca ccttgctcag ccccttttgcc cattttttaaa ttagattgcc ttttttatatt   133920
gagtttcagg agtcctttat atattctgaa taaatgtccc ttatcaaatt atatttatttc   133980
caggtatttt cttcattctg tgagttgtct ttcctctacc tttttaaaaa ggtgggttttt    134040
tgtttgtttg tttgtttgtt tttttaagat aaggtctcat tctgctgccc aggctggagt     134100
gcagtggcac aatcacagct cactgccacc tcaacttctcct gggccgaagt gatcctctta    134160
cttcagcctc ctgaatagct agggcatag atacacacta tcacacccag ctttttttttt    134220
ctgtttgtag agacagatct tactgtgttg cccaagttgg tctcaaactc taggctcaaa     134280
gtgattctcc cacctctgcc tcccagagtg ctgggattac aggtgtgagc cacacgcaac    134340
ctgtcttttn actattaata gtgtcttcct gcttcagcct cccgagtagc tgggattaca    134400
ggcacccac accatgcctg gctaatttt ttgcattttt agtagagaca gtgtttcacc     134460
atgttcaccc ggctggtctt gaactcctga cctcaggtga ttcacctgcc atggcctccc     134520
aaagtgctgg gattacaggc gtgagccact gcacccggcc aaaatattgc cttctaaca     134580
gtattgtctt ctaatttgtg aacatggatg tatcttcatg tatttatgtg ttcttttcatt     134640
tcagcagaat tttgtagttt tcagagtaga agcctttcac ctccttgggt catttattcc    134700
tatgttttaa gttcttttcg attccattat aaatagaatt gttttcttaa tttcatttttc     134760
agattgtttg atgagagagc atagaaaatac aagtgatttt tacatgttga tcttgcaact    134820
tcaactttga taaatctgat tgttagtct aatagttttc ttgtggattc ttagacttt      134880
tcaatatata agatcatgtc atttatggat agagatagtt ttttttctgg ctagaactta     134940
cagagcaatg atgagtagaa gtggcagaag caaaaatctt tgtcttgttt ctccaataga    135000
agggaaagct ttcagtttca tcatttaata tgatgttagg tgtgggtttt caataaatgc    135060
ctttttttcag attcaggaat ttccctatca ttcctgattt tttaaggcttt tttctatttt     135120
ttaaatcatg aaagggtgtt gaatattgtc atgttctttc tgtatcagta taaatgatcc    135180
tatgatttt gggttttatt ctgttgatgt gaaatattaa ttgatttcca gatgttaaac     135240
caaccttgca tacctgagat gaatctcact tggtcatggt gtataatctt ttcaatatgc    135300
```

88

```
tgctggattc catttactgg tattttgttg aagattttgt atctgaacgc ttaagataac   135360
atttacactc tatcagaaat gaattgacca taaatgtgag agtgtatttg tgggttcttg   135420
attctcttcc attccaaaga tagacataca tccgtctgta tgtctgtctt tatgccagta   135480
ccatactctc ttgattacta ttgctttgta ataagttttg aaatcagaaa gtataaatga   135540
gattttggta tctgagtaac agtcctcata gaattagttg ggaaatattc cctctttatt   135600
ctggtccctc tttctttttt gtttaactgt gtatcttgga gattgttcct tctgaaacaa   135660
tgagagccgc tttccctacc ctcccacccc tgctatagag aggtctataa gtgtctgttc   135720
aattatttta tttacttaac ctattactta gtcggggaca ttaagcttgt ttatgtcttt   135780
tattttaaac aatgctgcag tgaataatct tgtatataag tcattttcca tcaatataag   135840
tctctctgta actgaatttt tagaagtgga atttctaggt caacctatgg ctctgtattt   135900
cacaaaaata ccaattctgg tttttcttgt ggaggtgggg agtaggaggt agaatgctgg   135960
aggagaactt gctgtactca gctggctagt cattttagaa aggtttcctt agcttctttt   136020
tgtcatatgg cctcaccaag aatcaaaaac attcctattt accctgtaaa catggggctt   136080
tactacccaa gatacatatt tctggatgta tgacagcttt tcatattgaa gaaataatgc   136140
tgtgagtaca gcacatttgt tggaacttag gtcgttaaga atgtcttata aattcataca   136200
ttatacattt tattttattt tattttttag tttttgatac agagtctttcc tctgtcgccc   136260
aggccagcgt gcagtggtac aatcttggct cactgcgacc tccatctcct gggctcaagt   136320
gattctcatg tctcagcctc cagagtagct atggttacag gcatgcacca ccatgcccgg   136380
ctaatttttt tattttagt agaaactggg tttcaccata ttgaccatgc tggcctcgaa   136440
ctcttggcct caagtgatcg gcctgcctca gcctcccaaa gtgctgggat ccttgtattg   136500
ggtaaaagat gaatattgag ggctgcatgg tggctcatac ctgtaatccc agcactttct   136560
gagactgagg tgggaggat cctggagccc aggagggtga ggctgcagtg agttgtgatc   136620
gcgccattgc acttcaacct aggagatatta ggcttcagtc actgtgcccg gcatgtacat   136680
tttaatattg tgctttcctc ttttagctat agtatgaggt tacatttcag agtcattgtt   136740
gttaagcatc ttaatagtga tgaggttgag tgaaagttac ttctatttca aacactgaag   136800
aaaattttgt acaaatctgt cacattccaa gcccaggact gattgtttca tatacttcta   136860
attttacaat ttctattgta gtccagtgtg aaaaaagcca gtattaaaat actgaaaaat   136920
tttgatgaag cgataattgt ggatgcggca agtctggatc cagaatcttt atatcaacgg   136980
acatatgccg ggtaagctta gctcatgcct agaattttta caagtgtaaa taacttttgca   137040
tcttttaaat tttttaatta aattttacat ttttttctaa tctattatta tatgcccaga   137100
actttcactt agagtgtgca gtataatgtg gtggttaagt ataaaggctc tggagtgact   137160
tcctgggttt taatcttggc tctgccattt attggcagcc gctaacctct tggtatctca   137220
gtttcttcat ctgtaaaatg agataaataa agtgaaaaga tgccaacatc attttactctg   137280
ggctgcataa ctgatacttg gaaaaagtat tccttttgagt ttaagaatta agttggttat   137340
tcattttagc ttgtaataaa aagtagtga ttcataggat atgccactta ctgaagatta   137400
ccacagatcc aatcataaaa tcactttctc ttccctaaag atagcttgat taacatgtaa   137460
aggtgtgtaa aggcttgatt acactacccct gatccgtacc ccagttccca gcagcaccat   137520
gaaaaaggga tttcaacata ttaattacct ttcagtagaa agtaacagtg gtaggccagg   137580
cgcagtggct cacacctgta atcccagcac ttttgggaggc cgaggtgggc ggatcacgag   137640
gtcaggagat tgagaccatc ctggctaaca cgatgaaacc ccgtctctac taaaaataca   137700
aaaaattagc cgggcatggt ggcaggcacc tgtagtccca gctactcggg aggctgagac   137760
aggagaatgg cgtgagcccg ggaggcggag cttgcagtga gcttagattg tgccactgca   137820
ctccagcctg cgcagtggag cgagactctt gtctcaaaaa aaaagaaagt aacagtggta   137880
ttgggagact gaggagccta gaaagtactt gaaggaagta aaaggtttgt ttgaccacat   137940
tgtatttgga aagccagctt tttcagtcgt gtcagctttg tgtagtgatt tttagttctt   138000
cttttagaaa ataaccgaca ggtggctgca cggtggctca cgcctgtaat cccaccactt   138060
tgggaggccg agacgggcgg attacctgat tcaggagtt cgagaccagc ctgggcaaca   138120
tggtgaaacc ccgtctctac taaaatacaa aaagttagcc gggcgtggtg gcgtgtgcct   138180
gtagtcccag ctactccgga ggctgaggca ggagaattgc ttgaacccgg gaggcggagg   138240
ttgcagtgag ccaagatcac accattgcac tgcagccttgt gcgacagagt aagactctgt   138300
ctcaaaaaat aataataaaa taaaaagaa tggacagtaa acctaaatga gttcattccc   138360
aaagatgatg ttattcttaa gggatggttc atttatttaa gaccttacat aaagtctatc   138420
aattgcgtga tttttcactt ctgtaattgt gtgtatgtat aatgtaaata tatatgtttt   138480
tgttttgttt tggttttttg agacggagtc tcgctctgtt gtccaggctg ccgaggcagtg   138540
gtgcaatctc agctctctgc aacctctgcc tcccaggttc aagcgtttct tctgcctcat   138600
cctcccaagt agctgggact acaggcacgt gccaccacgc ccggctaatt ttttgtattt   138660
ttagtagaga tgggggtttca ccgtgttagc caggatggtc tcaatctcct gacctcgtga   138720
tccacccgcc ttggcttccc aaagtgttgc tattacaggc atgagccacc acaccagca   138780
tgtattttt aaatgtataa aatgaagcag aaaagagaaa tgataatttt tcttcatctt   138840
gaaagattat cttcaccagg cgcagtggct cacacttgta atcccagcac ttttgggaggc   138900
ctcggcaggc ggctcacttg agttcgaaac cagcctggcc gacatggtga aactccgtct   138960
ctactaaaaa taaataaaaaa aagatggttt tttagttta tttagtttta tgattttagc   139020
atcttttctga aatttttctc aaggcaagta aatttgtatc agttggtata ttggtaccca   139080
tctatgaaat aacttattag gaagatatct ctaaaataag atcacttttgc ctaaaataaa   139140
ctgatatatt gatgttcaca gaattttttct tttaaccgac ttgataaatg cattattctt   139200
gacgtcaagt gatccacctt cctcagcctc ccaaagtgct gggattacac acatgagcca   139260
ccgcacctgg cattattctt ataaaaggtt aaatttctag ttaagtttaa tgtcctcttt   139320
```

89

```
gttcatgtac cattgcttat tttcttccct tcctactcac agtaatcatt cttatggtat   139380
gcacttttgt ttgcttattt ttatgtaatt gatattacgc tccattctgt acgttgtcat   139440
ttcattcaca gtgagttttg gacattccta tgttcatcta tacagactta cttcatttta   139500
actacactgt agtattccgt atgtaatatt tactataact catcactgta gcagagcatc   139560
tcatagtgta tgtattactg ttttgccatt ttggtatcaa tgagtattta agtcatttgc   139620
agtttttccc tcttataccc agtattacag aggatctctt tttatatgct tctttgtacc   139680
aagaggcaga ttaaaaaatt tttttttgaa aaatttttg aaaaaaaatg aaatgaagtc   139740
tcactatgtt gcccaggctg gtctcaaact cctaggctca agcaatcctt ccatcttggc   139800
ctcccaaagt gctggggtta caggcatgag ccaccatgcc tggcctacat tttaaatttt   139860
gatagctctt acaatttact ttgtaaagta tctgcatcat tttatgttct caccagtctt   139920
taataagaat acttcatct tttggctgga cacagtggct cacgcctgta atcccagcac   139980
tttgggaggc cgaggcgggc agatcaagag atcgagacca ccctggccaa tatggtgaaa  140040
ccctgtctct actaaaaata caaaaattag ctgggcgtgg tggcgcaccc gtagtcccag  140100
ctactcgaga ggctgagaca ggagaatcac ttgaacccgg gaggtggagg ttgcagtgaa  140160
cttagatcac accactgcac tccagcctag caacagagtg agactctgtc tcaaaaaaaa  140220
aaaagaatac ttcagactta attttttttc cagtcttaag tgtttgctaa tgagattgag  140280
tttcttttgg tatgtctctt gattgttcag gtttttttct ttatgaattg actgttcatc  140340
tctttttcac attatttctg ttgggtgatt ttattagtga cttgttaaaa ttctgtatat  140400
tttttcagca tgacacttca ttattcaaaa aaaaaaaag attctctatg ttcctccgata  140460
ctaatcattg gttggtaata ccttaaaaat aagaccctta ctgtatttt tgcttttttt  140520
tttttttttt ttttttttt ttgagttag agtcttgctc tgttgcccag gctggagtgc  140580
aatggtgata tctcggctcct cagctcactg caactgcaac ctctacctcc ctgttcaag  140640
caattctcct gcctcagcct cccaagtagc tgggattaca ggcatccacc accacacca  140700
gctaattttt gtatttttag tagagacagg gtttcaccat gttggccagg ctggtcttcaa  140760
actactgggc tcaagtgatc cgcctgcctc ggcatcccaa agtactggga ttacaggcat  140820
gagccacagt gcctagccac tttttgcttt ttaactttgt tttatagtac tatagttta  140880
gtataaacag atgtatgtat acacacaact atggctttat aatatgttc agtcattgtt  140940
agagcaaggc ctactttttg ggtgcttctt ttacaaaatt gtctttggcta ttcttgtgcc  141000
ttttttctta tttgtgaatt ttagaattgt gaattacctg ttgactcacc atgttttgta  141060
aactgaggat tttgaatgga attgcactca attaaagatt atcttgcttt ctgtgcagca  141120
atgtttttatt tcaaataatc cctactttaa attacttagg atagctataa attgtgtttc  141180
tggctttctca gatttagatg aaacgcttta aattgattgt tttctcctaa atttaaaact  141240
gattgttaga agttaaagtc ttctgttcat tcttatttag gaagatgaca tttggaagag  141300
tcagtgactt ggggcaattc atccgagaat ctgagcctga acctgatgta aggaaatcaa  141360
aaggtttgtg gtgttttttat acttcatatt aagccttttac tcacattagt gattgactgt  141420
aagtcaaaga ccacttaagg tttaaactgt tatttttgta aagtaaccac tgtatctttc  141480
acccttgtgtt tatagtcaga agtaagtaca agggcttcct gtagtcacat ctttatgcaa  141540
tctcctctga atcaaaagtt agtgaacttg ctttgccact ccgaaaggca catgaatatg  141600
aaaaagcatt gtctattttc ttatttaatg gcaaaatacc cgacctaagt tggactttaat  141660
gtttgagacc gtttatttta taaattata tttttctct tttctttttt tttttgaga  141720
cagttcttgc tctgtcaccc agacccgggagt gcagtggtct gacgcacct cactgcaacc  141780
tctgcttcct aggttcaagc gattttcctg cctcatcctc ctgagtagct gggacttacaa  141840
gtgcgcacca ccacacctgg ctaatttttg tattttttagc agagatgagg ttcaccacg  141900
ttggccagc tggtctcata ctcctgacct caagcaatcc atccgcctctg gcttcccaaa  141960
gtgctgggat tacaagtgtg agccaccatg cctggcctta ttaaattatt tttattaaat  142020
ttcctcagaa ttgatgaaag taatgaaata taaaagtaat gaaatatatg tggaaaatag  142080
actggattaa gaaaatgtgg cacatataca ccatggatac tatgcagcca taaaaaagga  142140
tgagttcatg tcctttgtag ggacatggat gaagctggaa accatcattc tggcgaacaat  142200
gtctcaagga tagaaaacca aacaccgcat gctctcactc ataggtggga attgaacaat  142260
gagaacactt ggacacaggg tggggaacat cacacgctgg ggcctgtcgt agggggtggg  142320
gctgggggag gaatagcatt aggagatata cctaatataa atgacgagtt aatgggtgca  142380
gcacaccaac atggtacatg tatacatatg taacaaagct gcacgttgtg cacatgtacc  142440
ctagaactta aagtataata aatttaaaaa aaataaatat atgtggaaaa tattaatagg  142500
tcaaaattca aattgttcat ttaatcagaa gagtagttta gtcaaatcca agggttagac  142560
aacagaaatc ttttttgtca agtcattct ttgtgactga tttcattttc ttcctggttt  142620
acacaggaag atttcagaaa caaatgtgga tccgtgacag atggtatcta gaagtttttta  142680
gtttggttga attgacagta ttttattgag taaaagatac taattttgt aagaagaaaa  142740
attcaatttt gataagtatg tttaagatta agagctattg gccaggcgct gtggctcatg  142800
cctgtaatcc tagcactttg ggaagctgga gcaggtgggt cacgaggtca ataagattgag  142860
accatcctgg ccaacatggt gaaaccctgt ctctactaaa ttagccaggc gtggtggcac  142920
atgcctgtgc acccgcctcc gggtttaagc gatcctactg cctcaggctc ctgagtagct  142980
gggattacag gcgccatggc taattttagta atttttgta gagacgggt tttcactacat  143040
tggccaggct ggtctggtct caaactcctg acctcaggtg atctgcccgc cttagcctcc  143100
caaagtgctg ggattacagg catgattcac catgtctggc catttatctt atttttctttt  143160
ttttttttt ttttgtttga gacggagtct gctgtgtgtcg cccagactgc ggagtgcaatg  143220
gtgcgatctc agctcactgc aacctctgcc tcctggggctc aagcaattct cctgcctcag  143280
tcttccaagt agctgggatt acaggcgcgt gccaccacat ctagctaatt tttgtatttt  143340
```

91

```
acagagtgag accatgtgtc taaaaagtaa ataaaaatag tttctctttc atgactagaa    147420
tattacctct atgtgggcag ggagtttgtc tatactattt ggcactatat ttcctgattc    147480
tgaaattatg cctagcacat ggtaagtact ccttaaatat ttattgactg aattatttaa    147540
tacttaagaa tttcatttgg gattatctga gtggtaagat tacggattat atttatgtaa    147600
gaaaaaatca tttttttaaac ttggttgccc tttgccacac tgacatagac actaagtttt    147660
cttagccaga ttacttccga ggatactcac agaggccatt ctcttctcaa tccccaaata    147720
attgatattt cttagcactt tcaagctaat gcaattctta gatgatgtat ctgtgtatat    147780
catatcctca ttctacaaat gtagaaattg aagtctgagc acagtggctc tcacctgtaa    147840
tctcagcagt ttgggaggcc aaggcgagcg gatcactgag gacaagagtt aagaccagcc    147900
tggccaacat ggtaaagcct tgcctctatt aaaaatacaa caattagggc cgggcgtggt    147960
ggctcacgcc tataatccca gcacgttggg aggccaaggc aggcagatca cgaggtcagg    148020
agttcgagac catcctggct aacacagtga aaccccatct ctactaaaaa tacaaaaaat    148080
tagccaggca tggtggcacg cgcttgtagt cccagctatc gggaggctga ggcaggtgaa    148140
tcccttgaac ccgggaggcg gaggttgcaa tgagctgaga ttgcaccgct gaactccagc    148200
ctggtcaaca gagggagact ctgtctcaaa aaaaaaaaaa aaaaacaatt agccaggcgt    148260
ggtggcgggt acgagtacct gtaatcccag ctactaggga ggctgaggga ggagaatcac    148320
ttaaacccag gaggtggagt ttgcagcggg ctgataatgc accactacat tccagcctgg    148380
gcaacagagt gagactctgt cttaaaaaaa aaaaaaagaa agaaagaaat tgaggaatgt    148440
ggagattgtg gtctgtgatt tgttaggaat cacacagcag gttagtagca actacagggc    148500
tttggttcag aataccacct tgacaatggt ttgtttacag ttcggctccc cttcctctgc    148560
ctttctctcc ttccttattg agggcagctg gaaagaattt tcatcattta ctagcctata    148620
gctttaattt gagttttgaa accttgatca tagagcacag aggaaaagac tgagttttct    148680
ttttttgaga cagtcttgct ctatggccca ggctggagtg cagtgacacc atctcagctg    148740
gttgcaacct ctgcctccca ggttcaagca attctgcctc agcctctcga gtagctgaga    148800
ttacaggcac gtgtcaccac gcccagctaa ttttctgttt ttgtttcgtt ttgttttttt    148860
ctgagatgga gtcttgctct gtcacccagg ctggagtgca gtggtgcgat gttggctcac    148920
tcaaacctct gtctctggg ttcaagcaat tcttctgcct cagcctccc agtagctggg    148980
actacaggta cgtgccacca tccctagttc attttttgtat gtttagtaga gatggggttt    149040
cactatgttg accaggctgg tctcgaactc ctgtactcag gtgatctact cgtctcagtt    149100
tcccaaagtg ctgggattat tggcacacgc ctatttttgt attttttagta gagacggggt    149160
ttcaccatgt tggttagact ggtctcaaac ttctgacctc aagtgatttg ccccgcccag    149220
cctcccaaag tgctgggatt acaggcgtga gccaccgtgc ccagccaaga ttgagttttg    149280
aaaagagcct tctgagatta tgagaagggc aagcaagata acttaagaag ttacattaaa    149340
atcatctaag agacagtgta acaagaagga attgtaaaat gatgttatga gcacgtgccc    149400
aatgtagtgg caatcccttg tgcttcgata catgttgagg agacaaaact gtacttaaat    149460
tgataaatcc cttacatgtc attttaagga gcttagactg actcccatca tgtagacatc    149520
agagatttct tttttttttt ttttttttt ttttttgacag agttttgctc    149580
ttgttgccga ggctggagtg caatggcgtg atctcggctc accacaacct ccacctccca    149640
ggttcaagca attctcctgc ctcagcctcc cgagtagctg ggattacagc catgcaccac    149700
cacgcctggc taattttgta tttttagtag agacggggtt tctccatgtt gtggctggtc    149760
tcgaactcct gacctcaggt gatcctcccg cctcagccac ccaaagttct gaaattacag    149820
gcgtgagcca ccgcgcccag cccagagatt tctaaacaga gttctaacca gatgctttc    149880
cctgtcagta gaatgagaat gaattggagg tgggagagac tggcatgagg gacaccagtc    149940
agccagtgga attagctggt aatgttgata ggagaagaaa aagattcaaa gttaagtagt    150000
ggtagcaaga attagaggga aggtcggatt tatgatatgt ccaaggttga attctaaggt    150060
gaaatttggt ggcagatttc atgtgtaaat tgggaaggta gattgagttt tttaacatg    150120
ggttttctaa catgtcaata gagtgactct gcaggggggc ctgacgagag aacagtgcat    150180
ggggtgattc aacagccagt tgagccttca tgcagagcat ttaacactgt gactctgtag    150240
actctggttg gcagtaaaat ttcattaaac caatatttaa accttaggt aataataaaa    150300
attgagggaa aaggatccag gttttgtatt ttttatgaat tcagttattg aattaaacag    150360
gaccttgcct caagaaataa tctaccaaca attaacttgt tttaaagcaa agttaggaag    150420
tgagcatgtt caaattatta aataaaaaag taagctgtgt atttcattca tagaaataga    150480
ggctggccta cttcggatga ttctcagcat gtgattacag atgtgggctt atacatccta    150540
gggagttaag gcgtactctg gcttggatag agtagagctc tttgaaactc ttctctcacc    150600
cagctagttt atatagacta gagaactaga atgtagcagc atactctgtc ttagaagccc    150660
tttatatag gagctggtct ggaaggtttg aaaacataac aaatgtgttg gtgtctccca    150720
atgtattgct agattcttac ccaagagcat tatcctggtt agggtttggt ttggttttgt    150780
tttgtttttt aatgtttgcc acaaactaac actagatgtt agttctttca tcaagtgagg    150840
agagtagaag aaaaagtccag aactctgaaa caccttttca aaagtttttc aagccatgat    150900
gtttgcaagt taaatgctct gttatgtaag caatataatc agttttttatt aatgtaacat    150960
tccttagtgt tttggggtat cacacaaaaa agaatatcca tatctggaag caacagcttt    151020
taaataagag cattgtggtg gtggtggtga tagtggtttt tttttttttt tagttaagaa    151080
agtctcgctc tgttgcccag gttggagtgc agtggcacga tctcagctcg cttcaacctc    151140
tgctcccagg ttcaagcaat tcttctgcct cagcctcctg agtagctggg attataggca    151200
cctgctacca tgcctggctg attttttatta ttttagtaga cacaggtttc accatgttg    151260
ccaggctggt cttgaactct taacctcagg tgaatcaccc acctcggcct cccaaagtgc    151320
tggaattaca ggcatgaacc accatggcca gccaaataag agcattttta atgtaaaatt    151380
```

92

```
atgcatgaaa tgtacattca attttgtctt tgtttactag gatccatgtt ctcacaagct    151440
atgaagaaat gggtgcaagg aaatactgat gaggtaaatc ctacctttag gataaaaaga    151500
tttctgtttta taagtgccac cctcatgtaa gtgaggttta aattttcct tttctttagg    151560
tcccatgttt aagcagcatg gcacatttat gttctcttac ccagaatgta ccaagaaagg    151620
gtggtccctt cttaacatct aacaattgcc tggtagtagc agtgaaggta tcttcagtca    151680
gaggctagga ccactgaagg atatacatgc attcaagttt ccatcagcca gcaggcatca    151740
gtaatcagtg tgtagatcaa aagctcaaat gtttccttcc ccactggcag ttttacttca    151800
agtagtggag gcttgctttt ttaatagtta attaagtaca ttgagagatg ggaggtgaaa    151860
aaaggaaaat gttttatttt gaccatctaa tatgaaagta gttcggtgtt aggtatccag    151920
tagttgacac tggaagacag ggaatgacat gttaatattc atagccagag ggtggcccag    151980
gttttttcgt acatgggaat gaaattctta tccaaataag tagaaattat gtgcgtaagc    152040
catttgttaa gagcactgag tatgtgcatc tcgatccatc taatgaataa ccattatcac    152100
cagtttaaat tattttcttt aggcccagga agagctagct tggaagattg ctaaaatgat    152160
agtcagtgac attatgcagc aggctcagta tgatcaaccg ttagagaaat ctacaaaggt    152220
aaggatgact tcgtttttgtg taaactaaaa agtattattt tccaggtgta aaaataaaaa    152280
agaacataag gggttttcttt gcctttgaag gattaactgc tgtggggatt accttcttat    152340
cataagcaac tagaaaattg acaaactaaa tgaaacaact gtttgcatat attggacaat    152400
gggcaataca gggaaaccat ggaaaccaaa cagagcccag tagtcttgct gaacgaaaga    152460
gttaaatatc aaagttcagg ccaggtgcag tggctcacgc ctgtaatccc agcactttgg    152520
gaggccaagg cgggtgaatc acttgaggtc aggagtttca acaccgcctg gccaacatgg    152580
tgaaacctg tcttagccgg gtgtggtggc aggcacctgt aatcccaact attttgggagg    152640
ctgaggcagg agaatcgctt gaaccaggga ggcggaggtt gcagtgagcc gagatcacac    152700
cactgcactc cagcctgggc gacgagcgaa acccccatttc aaaaaaaaaa tcaaagttca    152760
gagagctcaa tttgagtaga agttgtagga taaggtagca gaaaagagga agctgcccag    152820
aaagaaagcc gtagagatat ttagagagat tcccatggat ccttggccta ggagtgatct    152880
gtatatgtgt ggggtgaaaa cgcatgtgtc caggtagaga accccccaga aattagtagg    152940
ctgaatgatt gctggaacat agggctaaga aaagttcatg gccagaagga tctggccaga    153000
gtagagagac ttagtaatac acaaggcatt gggtagtgtc ttacagagg ttatgccttta    153060
ctactgaaga taaattagtc ctagagtaca agcacctgaa ccaagtttca aagcaaattt    153120
ttaaaggggtc aaattaccta acaactgcat gccaaaacaa aggcctaacc ctctttacag    153180
taacacaaca aaattcagca cttcacagtg taaagttaga atgtctgacg tccaggctgg    153240
gcgcagtggc tcatgcctgt aatcccagca ctttgggagg ccgaggcagg tagatgacct    153300
gaggtcagga gttcaagacc agcctggcta acatggtgca accccgtctc tattaaaaat    153360
acaaaaactt agccaggcat ggtggcggc acctgtgatc ccggctactt gggaggctga    153420
ggcaggagaa ttgcctgaac ccaggaggtg aaggttgcag tgagccgaga tcgcaccact    153480
gcactctggt ctgggcaaaa agagcaaaac tcaggctcaa aaaaaaaaaa gaatgtctga    153540
cgtcaatcac aaattaccaa gcatgacatg aagttgacct ataaccagga gaaaactcaa    153600
tctatagaaa cagacccaga tgtgagaaag atgatgaatt tagcagacaa agaccatcaa    153660
gtggctattt taaatattaa aaatatgttc aagtggccag gtgcagtggc tcatgcctgt    153720
aatcccagca ctttgggagg ccaaggtggg taggagttca agaccagcct ggccaatatg    153780
gtgaaacccc ttctctacta aaaatacaaa aaaattagct gggcatggtg gcaggtgcct    153840
atagtcccag ctatatggga ggctgaggca caagaatcac ttgaacccgg gaggtggagg    153900
ttgaggttgc agtaagccga gattgtgcca cttgtactcc agcctggaca acagagtgag    153960
actctgtctc aaaaaaaaaa aaaaaaaagt taaagaaaac aagagtataa tgagaaaaat    154020
gcaaaatagt tttaaaagaa ccaaatggaa tttctaaaa taaaaaatac cagaaatggg    154080
ggccgggcgt ggtagctcac gtctataatc ccagcacttt gtgggagtg aggcaggcag    154140
atcacctgag atcgtagtt caaggccagc ctgaccaaca tggagaaaac tcatctctac    154200
taaaaataag aaattagctg ggcgtggtgg cgcattgcct gtaatcccag cgactctcgga    154260
ggctgaggca ggagaattgc ttgaacccgg gaggcagagg ttgcggtgag ctgagattgc    154320
accagtgcac tccagcttgg gccacaagag tgaaactccg tctcaaaaaa aaacaaaaa    154380
aaaacagtag actcgaagaa ctagctgagt ttttcctttac tttaggcagt aagtgtgacc    154440
ttttgcaggt gactacttta gttcctcatg tcctcattag tagatcagag aaattcgaca    154500
ccaaaacccc aaaagaaaaa ccccttctaa tcctcattcc atgattttat gaatgcatga    154560
agtcctaggc ctgcgaagga atactcattc tctttatcct gtgtttgatac ctctctgctt    154620
caacctccaa ctcgacattt gcctatagga tgtacttgga cattcagcat aaactaccttc    154680
acaccattac tgaattgctt catgtgcaca tgtcccatgc cacaataccg gggacccttgt    154740
cttcctgtgat atttgtccgc agtgctgtga ctacaggagg gagtcagtga atgtctgcat    154800
gtgtgtcttt accatccctc ttgaatatgc tctagggttta attcctagaa gtaaagttcaa    154860
tctattgaaa attggcaata tttttcattc taatatcat tgccaacatg gggaagcaag    154920
tctggatgcc agtccttgtt atatgcccct tgggtaagtt acgtaacctc ttataagcttc    154980
tgttcactca tattttaaca aggaaaatta caatattta cctcacaaaa ttgtagtcag    155040
cttctggctg tcttaaactc tggtatatag taaacactaa gtgttggtgt cactccttgaa    155100
tttgtaataa taggtcactt gttagagaaa tgcacctac cattttcttt tctttttcttt    155160
tttcagttat gactcaaaac ttgagataaa ggaaatctgc ttgtgaaaaa taagagaact    155220
tttttccctt ggttggattc ttcaacacag ccaatgaaaa cagcactata tttctgatct    155280
gtcactgttg tttccaggag agaatgggag acaatcctag acttccacca taatgcagtt    155340
acctgtaggc ataattgatg cacatgatgt tcacacagtg agagtcttaa agatacaaaa    155400
```

```
tggtattgtt tacattacta gaaaattatt agttttccaa tggcaataac ccatttatga   155460
gagtgtttta gcctactgga atagacaggg accacatcct ctgggaagca gataagcata   155520
gaactgatac ttgatgcaca ctcgtagtgg taactcatcc ctaatcagca ttgtaaagca   155580
ggtgccagag gtggtttgct ttgtccttcc aaagcaggtg agtcagcccc accgagagcc   155640
aggcagcttt gagtggcagc gtggtgctag cagcttcagc ggaacagggt gagagttaat   155700
tatgcagtct tcttgacagc ggcattaatt tggaaggaaa ctgacaagtc atgggtcaag   155760
tttcagtgac ttcctccttc ctctgatggc agtatatagt tttcacattt taattcctcc   155820
tcctgagatg cactatactt aaaaccattc tctcccctgc taacagaagg gtgtgaatct   155880
ggtttacttt gagcattagg atttgcccct ttggaattct gcactccagt tacttaactt   155940
tcccttcaga atacatgtgg aaagaaagaa agaaatagcg atgactccac ttttgcccct   156000
gtggcacctt gaacaaagca gttcttccca aattatactt tttttttttt taaataaggt   156060
gagcaggatg actggggaga gagaaacatt tgactttgac tgcctccccc attctttgct   156120
gtgagctgaa aagtgtgcag ttggtcgtct ttcttctcct ttctttagga tagtaagaga   156180
ctcactcact gcacttctgc tcagttggct tctgcatcgg gatcacacag ccatcagcag   156240
gactgcccag ttggtgagca cactccattg accacgcggc gccagcgctt cctcaatgca   156300
catgattgag aggaaagaaa gttctcttag atgttactgc ttttgctcag acttttgcaaa   156360
aaaaaaaata tatatatata atttattcca ttgcaatatt tgattgtata gaggcacact   156420
aagtcttgaa tgaacagaga aaaaaggctt tgtgtaagtt tttggtacta tgtacacact   156480
gttcatcgca cagaagaagc aaaaaggctt tgtgtaagtt tttggtacta tgtacacact   156540
ctgttattct tttaaagctg aagtattcat gtacttaaac catattatat taattgtgt   156600
ttgattttaa aatatatata tatgaattct atttaaaatt gtgtcaactt tctgctttca   156660
gggcatttat ggctcttctg ttgaaatata ttgatcttc caaatatttt catttgcttt   156720
ctaaaaaccc agaacatgag ccactactgg acttttgcctt gtgtttgaag tgtatggcat   156780
aaacccaagg tttttattag tcatctatgc tgtgattaat tcattttgtt cttttaacaa   156840
aatatttcca tccacttcac attgcttcaa tctttaacag aaaagcaata taaaggttat   156900
agaataaaat gtggtttttgg gcaactcttg ctgcctctgc atgttttgga ataacaattt   156960
ctacaagact ctaggctgtt taaactagtg ctttcagtta agataaattc taatcatttc   157020
tttgtatata catttttgtgc ttctgagcta gagatgccaa gtagttgtaa actgcttata   157080
aagagaatag cagcaaattt gagactcggc tactttttc tgccccacct gctttgagac   157140
acagaagcgg agtgtggccc gaaattatta gccagattta atatttgatc taaagtaggt   157200
ccttgtactc attttaaagt tggaatttga ttcctccaac attgagcacc caccatgttc   157260
caggctctgt gcattgtgcc cacaaaataa gattccctgg tggagttttt atgggttcaa   157320
ataatcagtt gaacaccctt catctttatc atgttgttga cattgacaca aattgtttaa   157380
aaagaaaaga tattagagag aaagtggtac ctttgtaact tgatgtgtct tcatcattcg   157440
gtaagattttg atgaaagtaa aaagcaaatg tcagccaaat ccagtgaaca gcaataaaac   157500
agggagtaac tttttataac tttttctact tggatttcaa cattcagtag agcttttcga   157560
aatgtaagta gtttcacgta ctggaggttt gactagttca gtaggaattt ggagggagaag   157620
gtcattctga attgtaacaa agtacaaact tctttgctgt tttatttaag tactgagagc   157680
taagcaccta atgaagtgac tgacctctcc ccagtgacag tgtttgggta cctgcctgcc   157740
ttcaggagtg gggtttatgt ttctacacag tgacctttc tctcgcccctc tcctccctct   157800
tgcccacaca ccagttgatt ggacctgggt tgaactcctg atccagacag gcccaagaca   157860
gttcttaatg ttaagaattt tggggccggg cacggtggct catgcctgta attgcaacac   157920
tttgggaggc cgagacaggc ggatcacttg aggtcaggag ttcgaggcca gcctggccaa   157980
catggtgaaa ccctgtcttt actaaaaata caaaaattag ctgggcatgg tggcgcacgc   158040
ctgtaatccc agctacgtgg gtggctgaga cagggggaatc gcttgaacct ggaggcggag   158100
gttgtgcaat gagccgagac cgtgtcactg cattccagcc tgggtgacag aggagagctc   158160
tgtctccaaa aataaaaata agaaaaagaa ttttgggcta ggtgcagtgg ctcacgcctg   158220
taattacagc attttggaag gcccaagatg ggcagatcac ttgaggacag gagttcgaaa   158280
ccagcctgga caacatggtg aaactccatc tctactaaaa agacaaaagt tagccagatg   158340
tggtgatggg cacctataat cctagctcct cgggaggctg gggcaggaga atcacttgaa   158400
cccaggaagc agagattgca gtgagccaag atcacatctc tgcactccag cctgggcaac   158460
agagcaagac tctgtctcaa aaaaaaaaga attttggccag gcgcagtggt tcacctgcgt   158520
aatcccagca ctttgggagg ccaaggcagg cagatcacga ggtcaggaga tcgagattgt   158580
cctggctaac atggtgaaac cctgtctcta ctaaaaatac aaaacattag ccgggtgtgg   158640
tggtgggcac ctgtagtccc agctactagg gaggctgagg cagaggaagg atgtgaaccc   158700
aggaggcgga gcttgcagta agccaagatc gtgccactgc actacagcct gggtgacaga   158760
gtgagactcc gtctcaaaaa aaaaaagaat ttgggcgggg tgcggtggca catgcctgta   158820
gtcccagcac tttgggaggc caaagtgggc ggatcacctg aggtcaggag ttcaagacca   158880
gtccggccaa tatggcgaaa ccctgtctct tactaaaaaa aatacaaaaa ttagccaggt   158940
gtggtggcgg gcacctgggg aggctgagac agggagaaat gcttgaaccg gggaggcgga   159000
ggttgcagta agccaagatc gtgccactgc actccagacc aagtctttt ctcaaaaaaa   159060
aaaaaaaaag aattttgcat ggggaaggag agatactgat caccatctgg aatggtgcct   159120
ggatgtggca cttacaaat caggagccag cactgcatgg acaaacagaa gcatgtgggc   159180
ctgatagaz aggtaccttg ataaccctga agacatcctt ggtttctgca tctattcctg   159240
catccttgca ttggactaca ttaatctgtc agttatcctt ataatgattt ttgatttttt   159300
tttttttgga tggagtttcg ctcttgttgc ccaggctgga gtgcaatggc acgatctcgg   159360
ctcaccacaa cctccacctc ccaggttcaa gtgattctgc tgcctcagcc tcctgagtaa   159420
```

                                              94

```
ctgggattac aggcatgcgc caccacacct ggctaatttt gtatttttag tagagacggg    159480
gtttctccat gttggtcagg ctggtctcga actcccaacc tcaggtgatc accctgtctc    159540
ggcctcccaa agtgctggga ttacaggcgt aagccatggt acccggtctg tttttttgatt   159600
ttttgaaacc agtctgaagt gagttttttt aattacgtga aaggagtttg gctaaaatac    159660
tgccatactg ccctaatgcc taatgattat gtattctcag catgtctgca aagtactgct    159720
gatttctgga gaataatttt tctttagtaa acttcactta agtcgtcatg ttgctatctct   159780
caaaatggta tcctaaccta atggagctaa aagacacccc ttgttttttat aacaagcagt    159840
tactgaggcc caggaagggg agaagtccct ggcttgtgag atgatcacca ttagaactca    159900
ggctgggcc agtgccttt catgcttctc agatccttcc aaagaataat gaagattata       159960
accgctttta gcaattgtaa taaacccaga aatagaaagc tttttggtta gagtactggt     160020
agaagtttgg cgggagagat aattttttaca aaatttgtaa atacctgcca attctatata    160080
ctaggcaagg tctctggcct tgtaaaaccc ctcaaggtta caactttggt ggcccacact     160140
aatagttacc cactgaggcc ctctccgggt gaacattgag cactagagga agccctctg      160200
cttgggcagg actgggcgtg gtgcagagta ggagcggtga tactgtggat tctgggcagg     160260
tggagatggc cagtgatgtc caataaagga cactggaggg agcagtgtga gtaaaggccc     160320
tgagggcatt catgttcagg gagggttgct gcccactggc ttgcttggca cacaggagag     160380
tgggtattcc tgccttagta acttttatgta aacaagtatt tcctcagtct gttcctctca    160440
aactgcctgc tctggcacat tcagaatgtc acagaactca cctggatgca ttcagcccct     160500
tgcctaaagg tgacagtgca tctccttccc caccccaccc ctcataccac tgaagcacct     160560
gtcagactgg cccagtctgt gggcaaggag cctagagagg gcttagtttc agcttgaaag     160620
gagctgggat ttaccaagaa gcaaatgaga gacgaggatt gcaacaactg tgccatttcc    160680
ccagcttcag ctgactcctg tatattgact gtgccttcag actcatccgt aagtgacccc    160740
aggctggcct ctcccacatc acagtaagaa ttccacacac catacaactt gagacagagg     160800
tccagctgaa ggaagcccca cacttctttc aagtttttct tagtcttctc ttcttggcaa    160860
agagtacctt ttgtttcttc taattatgta actattggtt tagtaaatat tcacccattc    160920
agtcaccctg taagtggcag gcactgttta cagggacaca ggaaggaata aaaacttgca    160980
ggcaccttgg agcttgcatt ctattgaaga ggtaatggaa gttgggatag cagctaaact    161040
atgctggtat tggccaggcg cagtggctca cacctgtaat cccagcactt tggaggccaa     161100
ggtgggcaga tcatgaagtc aggagatcga gaccatcctg gctaacatgg tgaaacccg      161160
tctctactaa aagtaaaaaa aaaaattagc caggtgtggt ggcgggcgcc tgtagtccca    161220
gctacttggg aggctgaggc aggagaatgg tgtgaaccca ggaggcgaag attgcagtga    161280
gccgagatgg caccactgca ctccagcctg ggtgacagag cgagactctg tctcagaaaa    161340
aaaaaatatg ctggtagttt tgattcaaga tggcctttgg agcccatgat ttaggtctcg     161400
tacccaccaa ggtctactgg aaaacatcag gctctcctgc tatagaccca tagggagagc    161460
tgcagccgag aggggagct gaagagaagt gcccctttctg tgtcctgtca gcctcatcct     161520
tccgcaagga ccagttgctg tgccactcca ttcacttgct gcaagactgg aggtttttcc   161580
tcaggtgttg agcacctggt ttacaagatg tcagcatctt gatgcctgag accatcaagg     161640
caagtctctg aacagggctt accttagagt aaggcttaga agaggccgta aagtcagtct     161700
cagctccgtg gctctgcaga gctttgggac atgtgaattc ttaaaaacaa gactattgta     161760
cagttactat atgcatgcag tataaaatta taaccttgga aaatcctagc tagctgttga    161820
gctaattcca taaagtaatc agctcctgag ttctgcagtg gtaataataa tcagcataat     161880
gagtaaacac tgtgtgtgcc aggcagcgtc tcatttgatc cttgtgataa tcttgtaagt     161940
actgattttc tcccttcttt aaacaaagtt ttttttttttt ttttagagag ggtctcacta    162000
tgttgcccag gctagtcttg aattc                                           162025

<210> 37
<211> 1350
<212> DNA
<213> Homo Sapien

<220>
<221> CDS
<222> (213)...(920)

<300>
<308> GenBank AJ242973
<309> 1999-10-26

<400> 37
gcggccgcgt cgacgtgaca gccggtacgc ccgggtttgg gcaacctcga ttacggcgg      60
cctccaggcc cgccagcagc gccccgcgc gccgccgc gccctgccg cccccggtt          120
ccggccggag accccactct ctgccgttcc cgcgtgcggct ccgctgccgag tagcgccgtc   180
cccggggacc acccttcggc tggcgccctc cc atg ctc tcg gcc acc cgg agg       233
                                    Met Leu Ser Ala Thr Arg Arg
                                     1               5

gct tgc cag ctc ctc ctc ctc cac agc ctc ttt ccc gtc ccg agg atg      281
```

95

```
Ala Cys Gln Leu Leu Leu Leu His Ser Leu Phe Pro Val Pro Arg Met
        10                      15                      20

ggc aac tcg gcc tcg aac atc gtc agc ccc cag gag gcc ttg ccg ggc    329
Gly Asn Ser Ala Ser Asn Ile Val Ser Pro Gln Glu Ala Leu Pro Gly
        25                      30                      35

cgg aag gaa cag acc cct gta gcg gcc aaa cat cat gtc aat ggc aac    377
Arg Lys Glu Gln Thr Pro Val Ala Ala Lys His His Val Asn Gly Asn
  40                      45                      50                55

aga aca gtc gaa cct ttc cca gag gga aca cag atg gct gta ttt gga    425
Arg Thr Val Glu Pro Phe Pro Glu Gly Thr Gln Met Ala Val Phe Gly
                    60                      65                      70

atg gga tgt ttc tgg gga gct gaa agg aaa ttc tgg gtc ttg aaa gga    473
Met Gly Cys Phe Trp Gly Ala Glu Arg Lys Phe Trp Val Leu Lys Gly
                75                      80                      85

gtg tat tca act caa gtt ggt ttt gca gga ggc tat act tca aat cct    521
Val Tyr Ser Thr Gln Val Gly Phe Ala Gly Gly Tyr Thr Ser Asn Pro
            90                      95                      100

act tat aaa gaa gtc tgc tca gaa aaa act ggc cat gca gaa gtc gtc    569
Thr Tyr Lys Glu Val Cys Ser Glu Lys Thr Gly His Ala Glu Val Val
        105                     110                     115

cga gtg gtg tac cag cca gaa cac atg agt ttt gag gaa ctg ctc aag    617
Arg Val Val Tyr Gln Pro Glu His Met Ser Phe Glu Glu Leu Leu Lys
  120                     125                     130                 135

gtc ttc tgg gag aat cac gac ccg acc caa ggt atg cgc cag ggg aac    665
Val Phe Trp Glu Asn His Asp Pro Thr Gln Gly Met Arg Gln Gly Asn
                    140                     145                     150

gac cat ggc act cag tac cgc tcg gcc atc tac ccg acc tct gcc aag    713
Asp His Gly Thr Gln Tyr Arg Ser Ala Ile Tyr Pro Thr Ser Ala Lys
                155                     160                     165

caa atg gag gca gcc ctg agc tcc aaa gag aac tac caa aag gtt ctt    761
Gln Met Glu Ala Ala Leu Ser Ser Lys Glu Asn Tyr Gln Lys Val Leu
        170                     175                     180

tca gag cac ggc ttc ggc ccc atc act acc gac atc cgg gag gga cag    809
Ser Glu His Gly Phe Gly Pro Ile Thr Thr Asp Ile Arg Glu Gly Gln
        185                     190                     195

act ttc tac tat gcg gaa gac tac cac cag cag tac ctg agc aag aac    857
Thr Phe Tyr Tyr Ala Glu Asp Tyr His Gln Gln Tyr Leu Ser Lys Asn
200                     205                     210                 215

ccc aat ggc tac tgc ggc ctt ggg ggc acc ggc gtg tcc tgc cca gtg    905
Pro Asn Gly Tyr Cys Gly Leu Gly Gly Thr Gly Val Ser Cys Pro Val
                    220                     225                     230

ggt att aaa aaa taa ttgctcccca catggtgggc ctttgaggtt ccagtaaaaa    960
Gly Ile Lys Lys  *
                235

tgctttcaac aaattgggca atgcttgtgt gattcacaat cgtggcattt aaagtgcaca   1020
aagtacaaag gaatttatac agattgggtt taccgaagta taatctatag gaggcgcgat   1080
ggcaagttga taaaatgtga cttatctcct aataagttat ggtgggagtg gagctgtgcg   1140
gtttcctgtg tcttctgggg tctgagtgaa gatagcaggg atgctgtgtt caccttctt   1200
ggtagaagct aaggtgtgag ctgggaggtt gctggacagg atggggggacc ccagaagtcc   1260
tttatctgtg ctctctgccc gccagtgcct tacaatttgc aaacgtgtat agcctcagtg   1320
actcattcgc tgaaatcctt cgctttacca                                     1350
```

96

<210> 38
<211> 235
<212> PRT
<213> Homo Sapien

<400> 38
Met Leu Ser Ala Thr Arg Arg Ala Cys Gln Leu Leu Leu Leu His Ser
 1   5   10   15
Leu Phe Pro Val Pro Arg Met Gly Asn Ser Ala Ser Asn Ile Val Ser
                20   25   30
Pro Gln Glu Ala Leu Pro Gly Arg Lys Glu Gln Thr Pro Val Ala Ala
        35   40   45
Lys His His Val Asn Gly Asn Arg Thr Val Glu Pro Phe Pro Glu Gly
    50   55   60
Thr Gln Met Ala Val Phe Gly Met Gly Cys Phe Trp Gly Ala Glu Arg
65   70   75   80
Lys Phe Trp Val Leu Lys Gly Val Tyr Ser Thr Gln Val Gly Phe Ala
                85   90   95
Gly Gly Tyr Thr Ser Asn Pro Thr Tyr Lys Glu Val Cys Ser Glu Lys
            100   105   110
Thr Gly His Ala Glu Val Val Arg Val Val Tyr Gln Pro Glu His Met
        115   120   125
Ser Phe Glu Glu Leu Leu Lys Val Phe Trp Glu Asn His Asp Pro Thr
    130   135   140
Gln Gly Met Arg Gln Gly Asn Asp His Gly Thr Gln Tyr Arg Ser Ala
145   150   155   160
Ile Tyr Pro Thr Ser Ala Lys Gln Met Glu Ala Ala Leu Ser Ser Lys
                165   170   175
Glu Asn Tyr Gln Lys Val Leu Ser Glu His Gly Phe Gly Pro Ile Thr
            180   185   190
Thr Asp Ile Arg Glu Gly Gln Thr Phe Tyr Tyr Ala Glu Asp Tyr His
        195   200   205
Gln Gln Tyr Leu Ser Lys Asn Pro Asn Gly Tyr Cys Gly Leu Gly Gly
    210   215   220
Thr Gly Val Ser Cys Pro Val Gly Ile Lys Lys
225   230   235

<210> 39
<211> 481
<212> DNA
<213> Homo Sapien

<300>
<308> GenBank AW195104
<309> 1999-11-29

<400> 39
ggcattattg gactgtaggt ttttattaaa acaaacattt ctcatagctc taagcaaagc   60
attagaattc atcaagcgga ctcacatctt ttctctgcac agagaggggc tgaaaaggga   120
gagaaagtcc cttatgtatg tctagatttg gtaaagcgaa ggatttcagc gaatgagtca   180
ctgaggctat acacgtttgc aaattgtaag gcactggcgg gcagagagca cagataaagg   240
acttctgggg tcccccatcc tgtccagcaa cctcccagct cacaccttag cttctaccaa   300
gaagggtgaa cacagcatcc ctgctatctt cactcagacc ccagaaaacc cagggaaacc   360
cgacagctcc actcccacca taacttatta ggagataagt cacattttat caacttgcca   420
tcgcgcctcc tatagattat acttcggtaa acccaatctg tataaattcc tttgtacttt   480
g   481

<210> 40
<211> 390
<212> DNA
<213> Homo Sapien

<300>
<308> GenBank AW874187
<309> 2000-05-22

<400> 40

97

```
tttttttttat tggactgtag gttttatta aaacaaacat ttctcatagc tctaagcaaa        60
gcattagaat tcatcaagcg gactcacatc ttttctctgc acagagaggg ctgaaaaggg       120
agagaaagcc cctatgtat gtctagattt ggtaaagcga aggatttcag cgaatgagtc        180
actgaggcta tacacgtttg caaattgtaa ggcactggcg ggcagagagc acagataaag       240
gactttggg ggtcccccat tcctgtccag caacctccca gctcacacct tagcttctac        300
caagaagggg tgaacacagc atccctgcta tcttcactca gacccccaga agacacagga       360
aaccgcacag ctccactccc accataactt                                        390
```

<210> 41
<211> 43
<212> DNA
<213> Artificial Sequence

<220>
<223> Oligonucleotide Primer

<400> 41
agcggataac aatttcacac agggagctag cttggaagat tgc   43

<210> 42
<211> 22
<212> DNA
<213> Artificial Sequence

<220>
<223> Oligonucleotide Primer

<400> 42
gtccaatata tgcaaacagt tg   22

<210> 43
<211> 23
<212> DNA
<213> Artificial Sequence

<220>
<223> Oligonucleotide Primer

<400> 43
agcggataac aatttcacac agg   23

<210> 44
<211> 18
<212> DNA
<213> Artificial Sequence

<220>
<223> Oligonucleotide Primer

<400> 44
actgagcctg ctgcataa   18

<210> 45
<211> 21
<212> DNA
<213> Artificial Sequence

<220>
<223> Oligonucleotide Primer

<400> 45
tctcaatcat gtgcattgag g   21

<210> 46
<211> 43
<212> DNA
<213> Artificial Sequence

<220>
<223> Oligonucleotide Primer

<400> 46
agcggataac aatttcacac aggatcaca cagccatcag cag   43

<210> 47
<211> 23
<212> DNA
<213> oligonucleotide primer

<400> 47
agcggataac aatttcacac agg   23

<210> 48
<211> 18
<212> DNA
<213> Oligonucleotide primer

<400> 48
ctggcgccac gtggtcaa   18

<210> 49
<211> 20
<212> DNA
<213> Artificial Sequence

<220>
<223> Oligonucleotide Primer

<400> 49
tttctctgca cagagagggc   20

<210> 50
<211> 44
<212> DNA
<213> Artificial Sequence

<220>
<223> Oligonucleotide Primer

<400> 50
agcggataac aatttcacac agggctgaaa tccttcgctt tacc   44

<210> 51
<211> 23
<212> DNA
<213> Artificial Sequence

<220>
<223> Oligonucleotide Primer

<400> 51
agcggataac aatttcacac agg   23

<210> 52
<211> 18
<212> DNA
<213> Artificial Sequence

<220>
<223> Oligonucleotide Primer

<400> 52
ctgaaaaggg agagaaag   18

<210> 53

```
<211> 20
<212> DNA
<213> Artificial Sequence

<220>
<223> Oligonucleotide Primer

<400> 53
tcccaaagtg ctggaattac                                    20

<210> 54
<211> 22
<212> DNA           .
<213> Artificial Sequence

<220>
<223> Oligonucleotide Primer

<400> 54
gtccaatata tgcaaacagt tg                                  22

<210> 55
<211> 20
<212> DNA
<213> Artificial Sequence

<220>
<223> Oligonucleotide Primer

<400> 55
cccacagcag ttaatccttc                                    20

<210> 56
<211> 18
<212> DNA
<213> Artificial Sequence

<220>
<223> Oligonucleotide primer

<400> 56
gcgctcctgt cggtgcca                                      18

<210> 57
<211> 18
<212> DNA
<213> Artificial Sequence

<220>
<223> Oligonucleotide primer

<400> 57
gcctgactgg tggggccc                                      18

<210> 58
<211> 15
<212> DNA
<213> Artificial Sequence

<220>
<223> Oligonucleotide primer

<400> 58
catgcatgca cggtc                                         15

<210> 59
<211> 30
```

<212> DNA
<213> Artificial Sequence

<220>
<223> Oligonucleotide primer

<400> 59
cagagagtac ccctcgaccg tgcatgcatg   30

<210> 60
<211> 15
<212> DNA
<213> Artificial Sequence

<220>
<223> Oligonucleotide primer

<400> 60
catgcatgca cggtt   15

<210> 61
<211> 30
<212> DNA
<213> Artificial Sequence

<220>
<223> Oligonucleotide primer

<400> 61
gtacgtacgt gccaactccc catgagagac   30

<210> 62
<211> 14
<212> DNA
<213> Artificial Sequence

<220>
<223> Oligonucleotide primer

<400> 62
catgcatgca cggt   14

<210> 63
<211> 18
<212> DNA
<213> Artificial Sequence

<220>
<223> Oligonucleotide primer

<400> 63
gcctgactgg tggggccc   18

<210> 64
<211> 26
<212> DNA
<213> Artificial Sequence

<220>
<223> Oligonucleotide primer

<400> 64
gtgctgcagg tgtaaacttg taccag   26

<210> 65
<211> 28
<212> DNA

<210> Artificial Sequence

<220>
<223> Oligonucleotide primer

<400> 65
cacggatccg gtagcagcgg tagagttg   28

<210> 66
<211> 19
<212> DNA
<213> Artificial Sequence

<220>
<223> Oligonucleotide primer

<400> 66
actgggcatg tggagacag   19

<210> 67
<211> 18
<212> DNA
<213> Artificial Sequence

<220>
<223> Oligonucleotide primer

<400> 67
gcactttctt gccatgag   18

<210> 68
<211> 14
<212> DNA
<213> Artificial Sequence

<220>
<223> Oligonucleotide primer

<400> 68
tcagtcacga cgtt   14

<210> 69
<211> 14
<212> DNA
<213> Artificial Sequence

<220>
<223> Oligonucleotide primer

<400> 69
cggataacaa tttc   14

<210> 70
<211> 37
<212> DNA
<213> Artificial Sequence

<220>
<223> Oligonucleotide primer

<400> 70
caatttcatc gctggatgca atctgggcta tgagatc   37

<210> 71
<211> 37
<212> DNA
<213> Artificial Sequence

<220>
<223> Oligonucleotide primer

<400> 71
caatttcaca cagcggatgc ttcttttggc tctgact   37

<210> 72
<211> 40
<212> DNA
<213> Artificial Sequence

<220>
<223> Oligonucleotide primer

<400> 72
tcagtcacga cgttggatgc caataaaagt gactctcagc   40

<210> 73
<211> 37
<212> DNA
<213> Artificial Sequence

<220>
<223> Oligonucleotide primer

<400> 73
cggataacaa tttcggatgc actgggagca ttgaggc   37

<210> 74
<211> 38
<212> DNA
<213> Artificial Sequence

<220>
<223> Oligonucleotide primer

<400> 74
tcagtcacga cgttggatga gcagatccct ggacaggc   38

<210> 75
<211> 38
<212> DNA
<213> Artificial Sequence

<220>
<223> Oligonucleotide primer

<400> 75
cggataacaa tttcggatgg acaaaatacc tgtattcc   38

<210> 76
<211> 36
<212> DNA
<213> Artificial Sequence

<220>
<223> Oligonucleotide primer

<400> 76
tcagtcacga cgttggatgc agagcagctc cgagtc   36

<210> 77
<211> 32
<212> DNA
<213> Artificial Sequence

<220>

103

<223> Oligonucleotide primer

<400> 77
cagcggtgat cattggatgc aggaagctct gg   32

<210> 78
<211> 38
<212> DNA
<213> Artificial Sequence

<220>
<223> Oligonucleotide primer

<400> 78
tcagtcacga cgttggatgc ccacatgcca cccactac   38

<210> 79
<211> 35
<212> DNA
<213> Artificial Sequence

<220>
<223> Oligonucleotide primer

<400> 79
cggataacaa tttcggatgc ccgtcaggta ccacg   35

<210> 80
<211> 37
<212> DNA
<213> Artificial Sequence

<220>
<223> Oligonucleotide primer

<400> 80
tcagtcacga cgttggatgc ccacagtgga gcttcag   37

<210> 81
<211> 22
<212> DNA
<213> Artificial Sequence

<220>
<223> Oligonucleotide primer

<400> 81
gctcatacct tgcaggatga cg   22

<210> 82
<211> 36
<212> DNA
<213> Artificial Sequence

<220>
<223> Oligonucleotide primer

<400> 82
tcagtcacga cgttggatga ccagctgttc gtgttc   36

<210> 83
<211> 34
<212> DNA
<213> Artificial Sequence

<220>
<223> Oligonucleotide primer

104

<400> 83
tacatggagt tcggggatgc acacggcgac tctc   34

<210> 84
<211> 40
<212> DNA
<213> Artificial Sequence

<220>
<223> Oligonucleotide primer

<400> 84
tcagtcacga cgttggatgg ggaagagcag agatatacgt   40

<210> 85
<211> 29
<212> DNA
<213> Artificial Sequence

<220>
<223> Oligonucleotide primer

<400> 85
gaggggctga tccaggatgg gtgctccac   29

<210> 86
<211> 30
<212> DNA
<213> Artificial Sequence

<220>
<223> Oligonucleotide primer

<400> 86
tgaagcactt gaaggatgag ggtgtctgcg   30

<210> 87
<211> 38
<212> DNA
<213> Artificial Sequence

<220>
<223> Oligonucleotide primer

<400> 87
cggataacaa tttcggatgc tgcgtgatga tgaaatcg   38

<210> 88
<211> 26
<212> DNA
<213> Artificial Sequence

<220>
<223> Oligonucleotide primer

<400> 88
gatgaagctc ccaggatgcc agaggc   26

<210> 89
<211> 27
<212> DNA
<213> Artificial Sequence

<220>
<223> Oligonucleotide primer

<400> 89

```
gccgccggtg taggatgctg ctggtgc                                              27

<210> 90
<211> 31
<212> DNA
<213> Artificial Sequence

<220>
<223> Oligonucleotide Template

<400> 90
cgcagggttt cctcgtcgca ctgggcatgt g                                         31

<210> 91
<211> 43
<212> DNA
<213> Artificial Sequence

<220>
<223> Biotinylatd primer

<400> 91
tgcttatccc tgtagctacc ctgtcttggc cttgcagatc caa                            43

<210> 92
<211> 42
<212> DNA
<213> Artificial Sequence

<220>
<223> Oligonucleotide primer

<400> 92
agcggataac aatttcacac aggccatcac accgcggtac tg                             42

<210> 93
<211> 44
<212> DNA
<213> Artificial Sequence

<220>
<223> Oligonucleotide primer

<400> 93
cccagtcacg acgttgtaaa acgtcttggc cttgcagatc caag                           44

<210> 94
<211> 42
<212> DNA
<213> Artificial Sequence

<220>
<223> Oligonucleotide primer

<400> 94
agcggataac aatttcacac aggccatcac accgcggtac tg                             42

<210> 95
<211> 20
<212> DNA
<213> Artificial Sequence

<220>
<223> Oligonucleotide primer

<400> 95
ctccagctgg gcaggagtgc                                                      20
```

```
<210> 96
<211> 17
<212> DNA
<213> Artificial Sequence

<220>
<223> Oligonucleotide primer

<400> 96
cacttcagtc gctccct                                                        17

<210> 97
<211> 23
<212> DNA
<213> Artificial Sequence

<220>
<223> Biotinylated primer

<400> 97
cccagtcacg acgttgtaaa acg                                                 23

<210> 98
<211> 100
<212> DNA
<213> Homo sapien

<400> 98
cctttgagaa agggctctgc ttgagttgta gaaagaaccg ctgcaacaat ctgggctatg        60
agatcaataa agtcagagcc aaaagaagca gcaaaatgta                              100

<210> 99
<211> 100
<212> DNA
<213> Homo sapien

<400> 99
cctttgagaa agggctctgc ttgagttgta gaaagaaccg ctgcaacaat ctgggctatg        60
agatcagtaa agtcagagcc aaaagaagca gcaaaatgta                              100

<210> 100
<211> 100
<212> DNA
<213> Homo sapien

<400> 100
gaattatttt tgtgtttcta aaactatggt tcccaataaa agtgactctc agcgagcctc        60
aatgctccca gtgctattca tgggcagctc tctgggctca                              100

<210> 101
<211> 100
<212> DNA
<213> Homo sapien

<400> 101
gaattatttt tgtgtttcta aaactatggt tcccaataaa agtgactctc agcaagcctc        60
aatgctccca gtgctattca tgggcagctc tctgggctca                              100

<210> 102
<211> 84
<212> DNA
<213> Homo sapien

<400> 102
taataggact acttctaatc tgtaagagca gatccctgga caggcgagga atacaggtat        60
tttgtccttg aagtaacctt tcag                                               84
```

<210> 103
<211> 84
<212> DNA
<213> Homo sapien

<400> 103
taataggact acttctaatc tgtaagagca gatccctgga caggcaagga atacaggtat   60
tttgtccttg aagtaacctt tcag   84

<210> 104
<211> 100
<212> DNA
<213> Homo sapien

<400> 104
ctcaccatgg gcatttgatt gcagagcagc tccgagtccg tccagagctt cctgcagtca   60
atgatcaccg ctgtgggcat ccctgaggtc atgtctcgta   100

<210> 105
<211> 100
<212> DNA
<213> Homo sapien

<400> 105
ctcaccatgg gcatttgatt gcagagcagc tccgagtcca tccagagctt cctgcagtca   60
atgatcaccg ctgtgggcat ccctgaggtc atgtctcgta   100

<210> 106
<211> 100
<212> DNA
<213> Homo sapien

<400> 106
agcaaggact cctgcaaggg ggacagtgga ggcccacatg ccacccacta ccagggcacg   60
tggtacctga cgggcatcgt cagctggggc cagggctgcg   100

<210> 107
<211> 100
<212> DNA
<213> Homo sapien

<400> 107
agcaaggact cctgcaaggg ggacagtgga ggcccacatg ccacccacta ccggggcacg   60
tggtacctga cgggcatcgt cagctggggc cagggctgcg   100

<210> 108
<211> 100
<212> DNA
<213> Hom sapien

<400> 108
caataactct aatgcagcgg aagatgacct gcccacagtg gagcttcagg gcgtggtgcc   60
ccggggcgtc aacctgcaag gtatgagcat accccccttc   100

<210> 109
<211> 100
<212> DNA
<213> Homo sapien

<400> 109
caataactct aatgcagcgg aagatgacct gcccacagtg gagcttcagg gcttggtgcc   60
ccggggcgtc aacctgcaag gtatgagcat accccccttc   100

<210> 110
<211> 100
<212> DNA
<213> Homo sapien

108

<400> 110
ttgaagcttt gggctacgtg gatgaccagc tgttcgtgtt ctatgatcat gagagtcgcc   60
gtgtggagcc ccgaactcca tgggtttcca gtagaatttc   100

<210> 111
<211> 100
<212> DNA
<213> Homo sapien

<400> 111
ttgaagcttt gggctacgtg gatgaccagc tgttcgtgtt ctatgatgat gagagtcgcc   60
gtgtggagcc ccgaactcca tgggtttcca gtagaatttc   100

<210> 112
<211> 100
<212> DNA
<213> Homo sapien

<400> 112
ggataacctt ggctgtaccc cctgggggaag agcagagata tacgtgccag gtggagcacc   60
caggcctgga tcagcccctc attgtgatct gggagccctc   100

<210> 113
<211> 100
<212> DNA
<213> Homo sapien

<400> 113
ggataacctt ggctgtaccc cctgggggaag agcagagata tacgtaccag gtggagcacc   60
caggcctgga tcagcccctc attgtgatct gggagccctc   100

<210> 114
<211> 80
<212> DNA
<213> Homo sapien

<400> 114
tgaagcactt gaaggagaag gtgtctgcgg gagccgattt catcatcacg cagcttttct   60
ttgaggctga cacattcttc   80

<210> 115
<211> 80
<212> DNA
<213> Homo sapien

<400> 115
tgaagcactt gaaggagaag gtgtctgcgg gagtcgattt catcatcacg cagcttttct   60
ttgaggctga cacattcttc   80

<210> 116
<211> 80
<212> DNA
<213> Homo sapien

<400> 116
tccagatgaa gctcccagaa tgccagaggc tgctccccgc gtggcccctg caccagcagc   60
tcctacaccg gcggcccctg   80

<210> 117
<211> 80
<212> DNA
<213> Homo sapien

<400> 117
tccagatgaa gctcccagaa tgccagaggc tgctcccccc gtggcccctg caccagcagc   60
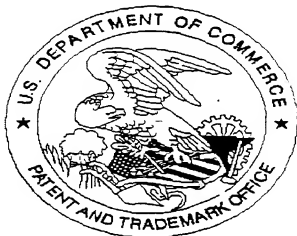tcctacaccg gcggcccctg   80

109

<210> 118
<211> 48
<212> DNA
<213> Artificial Sequence

<220>
<223> Hair pin structure

<400> 118
cagagagtac ccctcaaccg tgcatgcatg aaacatgcat gcacggtt   48

SCANNED. # _b_

Application deficiencies were found during scanning:

☐ Page(s)_____ of _____ were not present
for scanning.   (Document title)

☐ Page(s)_____ of _____ were not present
for scanning.   (Document title)

☑ Scanned copy is best available. *Drawings*